

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2004, 08:13:15 ; Search time 382.871 Seconds
(without alignments)
10794.564 Million cell updates/sec

Title: US-09-729-674-1_COPY_132_1265

Perfect score: 1134
Sequence: 1 gggcgcccaatacgaagc.....tgggaactggagagaccagc 1134

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2434939 seqs, 1822278265 residues

Total number of hits satisfying chosen parameters: 4869878

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1133.2	99.9	3871	9	US-09-729-674-1
2	850	75.0	1810	12	US-10-108-260A-2059
3	455.4	40.2	2347	10	US-09-562-832-113
4	455.4	40.2	2347	12	US-10-439-388-64
5	455.4	40.2	2443	13	US-10-094-749-724
6	196.6	17.3	398	12	US-10-242-535A-8221
7	154.8	13.7	2600	14	US-10-045-815-3
8	154.8	13.7	2829	14	US-10-045-815-1
9	153.8	13.6	892	13	US-10-027-632-161474
10	153.8	13.6	892	13	US-10-027-632-161475
11	153.8	13.6	892	13	US-10-027-632-161476
12	153.8	13.6	892	14	US-10-027-632-161474
13	153.8	13.6	892	14	US-10-027-632-161475
14	153.8	13.6	892	14	US-10-027-632-161476
15	147.4	13.0	2322	14	US-10-045-815-7

16	147.4	13.0	2416	14	US-10-045-815-5
17	145.4	12.8	564	13	US-10-027-632-114858
18	145.4	12.8	564	14	US-10-027-632-114858
19	138.4	11.3	3396	13	US-10-094-749-921
20	126	11.1	618	10	US-09-879-536-215
21	109.2	9.6	2665	14	US-10-071-766-86
22	107.2	9.5	226	12	US-10-242-535A-9365
23	103.4	9.1	150	12	US-10-242-535A-1775
24	97.8	8.6	2753	10	US-09-728-952-28
25	97.8	8.6	2753	10	US-09-728-952-51
26	87	7.7	349	9	US-09-925-303-395
27	73	6.4	247	12	US-10-242-535A-4919
28	69.6	6.1	181	13	US-10-023-386-16058
29	69.6	6.1	549	13	US-10-023-386-2358
30	62.4	5.5	3897	14	US-10-412-000-7
31	62.4	5.5	3897	14	US-10-021-657-7
32	55.8	4.9	2404	12	US-10-260-238-26
33	53.8	4.7	2305	12	US-10-094-886-23
34	53.4	4.7	1104	15	US-10-156-761-6371
35	53.4	4.7	9025608	15	US-10-156-761-1
36	53.2	4.7	1374	15	US-10-156-761-7192
37	52.6	4.6	1906	13	US-10-412-000-1
38	52.6	4.6	1906	14	US-10-021-657-1
39	52.2	4.6	966	15	US-10-156-761-3002
40	52.2	4.6	1413	15	US-10-156-761-4744
41	52.2	4.6	9025608	15	US-10-156-761-1
42	52	4.6	3624	11	US-09-988-462-6
43	51.6	4.6	461	13	US-10-029-386-20388
44	51.6	4.6	2793	15	US-10-156-761-2591
45	51.4	4.5	3275	13	US-10-027-632-114407

ALIGNMENTS

RESULT 1

US-09-729-674-1
; Sequence 1, Application US/09729674
; Patent No. US20010039335A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: LaVallie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Agostino, Michael J.
; APPLICANT: Steindinger II, Robert J.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fecthel, Kim
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 6055-64X
; CURRENT APPLICATION NUMBER: US/09/729,674
; CURRENT FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: 09/539,330
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 283
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3871
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-729-674-1

Query Match 99.9%; Score 1133.2; DB 9; Length 3871;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGCGGCCCAATACGAACGTACAGCTTCGCGCTCCACGGAGCGAGTGCGG 60

132 GGGCGCGCCCAATACGAACCTCAGCTTCGGAGCTTCCACGGGACGAGCTGATGCCG 191
61 CTCAGTCTGGCTACCGCGCACGCTCGAAGTACAGCGCGGAGCACTGGGCGGAGAGC 120
192 CTCAGTCTGGCTACCGCGCACGCTCGAAGTACAGCGCGGAGCACTGGGCGGAGAGC 251
121 STKGCTACCTGGAGATCAGCTCGGCTGCAAGCTTGTGGCGACGAGGCTTC 180
252 STKGCTACCTGGAGATCAGCTCGGCTGCAAGCTTGTGGCGACGAGGCTTC 311
181 TGCCACCGCAACTCAGCGCGCGCCGAGCCCGGAGCCCGCGGCTTCGCGAGCTAT 240
312 TGCCACCGCAACTCAGCGCGCGCCGAGCCCGGAGCCCGCGGCTTCGCGAGCTAT 371
241 CCCGAGTGGCCCTCTTGGGGGCTGCTGGCGCGCGGAGCTTCGCTCAAGCGGTGAAG 300
372 CCCGAGTGGCCCTCTTGGGGGCTGCTGGCGCGCGGAGCTTCGCTCAAGCGGTGAAG 431
301 CAGGCGCTGCGAGCTTCCGCGAGTCCGAGCCCGGAGCGGAGGTGCTGGCGGACTTCAG 360
432 CAGGCGCTGCGAGCTTCCGCGAGTCCGAGCCCGGAGCGGAGGTGCTGGCGGACTTCAG 491
361 CGCGCGAGCCCTACAAGTTCCTGAGTTCGCTTACTTCAAGGCAATATCTCCCAAA 420
492 CGCGCGAGCCCTACAAGTTCCTGAGTTCGCTTACTTCAAGGCAATATCTCCCAAA 551
421 GCCATCGCGCTGCTCACACCTTCTACTGAGCATCTGATGACGAATGATGAAGAGG 480
552 GCCATCGCGCTGCTCACACCTTCTACTGAGCATCTGATGACGAATGATGAAGAGG 611
481 AACATGCGCATATTAAGAGCTGCTGCTGCGGAGGACTACTTAAAGACCTGGAAC 671
612 AACATGCGCATATTAAGAGCTGCTGCTGCGGAGGACTACTTAAAGACCTGGAAC 731
541 AAGTCATATGAAGCTGTTATCGAGAGTGGGGCATCAAGCTGAGGAGTGAAGTGGAGA 600
672 AAGTCATATGAAGCTGTTATCGAGAGTGGGGCATCAAGCTGAGGAGTGAAGTGGAGA 731
601 ACATCCATCACAGATGAGGCTGCGCTTCCGAGCTTCTTCAAGGCTTTTACGAGTGT 660
732 ACATCCATCACAGATGAGGCTGCGCTTCCGAGCTTCTTCAAGGCTTTTACGAGTGT 791
661 CTCGAGCTGCGAGGTTCCAGGAGATCAAGGACTTCAAGGATTTCTACTTCCATA 720
792 CTCGAGCTGCGAGGTTCCAGGAGATCAAGGACTTCAAGGATTTCTACTTCCATA 851
721 GCAGATCATATGTAGAGTTCTGGAAATGCAAAATACAGTGTGAAGAACTCACCCCA 780
852 GCAGATCATATGTAGAGTTCTGGAAATGCAAAATACAGTGTGAAGAACTCACCCCA 911
781 GTTATAGGAGGCTATCCGTTGAGAAATTTGGGCTACCATGTATCATTTACGAGTTT 840
912 GTTATAGGAGGCTATCCGTTGAGAAATTTGGGCTACCATGTATCATTTACGAGTTT 971
841 GCCTATATATAGTTGAACACTGAGATGAGCCCGCTGCGAGTTCAGTCTGCTC 900
972 GCCTATATATAGTTGAACACTGAGATGAGCCCGCTGCGAGTTCAGTCTGCTC 1031
901 TTTGATCAGAAATGCAAGGTCATGACAGCAACCTGGTGTATTAACAGTACACAGGAGC 960
1032 TTTGATCAGAAATGCAAGGTCATGACAGCAACCTGGTGTATTAACAGTACACAGGAGC 1091
961 ACTTGGGGCTCTCGATGAGCACTTCCAGCCGAGACTGAGGAGTTCAGTCTTCTTAAT 1020
1092 ACTTGGGGCTCTCGATGAGCACTTCCAGCCGAGACTGAGGAGTTCAGTCTTCTTAAT 1151
1021 GTGACACACTCTCAGAGGAGCTGTATGACTTTGCTAAGGAAATATATATGATGATGAT 1080
1152 GTGACACACTCTCAGAGGAGCTGTATGACTTTGCTAAGGAAATATATATGATGATGAT 1211
1081 GAGGAGAAATTTGTGAATATGTGGATGACCTTTGGAACTGGAGGAGACCGAGC 1134

Db 1212 GAGGGAGAAGTTGTGGAATATGTGATGACCTCTTGGAACTGGAGGAGACCGAGC 1265
RESULT 2
US-10-108-260A-2058
; Sequence 2058, Application US/10108260A
; Publication No. US2004000560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US2004000560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2058
; LENGTH: 1810
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-2058

Query Match 75.0%; Score 850; DB 12; Length 1810;
Best Local Similarity 88.2%; Pred. No. 1.9e-243;
Matches 1000; Conservative 2; Mismatches 2; Indels 130; Gaps 2;
QY 1 GGGCGCGCCCAATACGAAGCTTACAGCTTCGCGAGCTTCCACGGGACGAGCTGATGCCG 60
Db 94 GGGCGCGCCCAATACGAAGCTTACAGCTTCGCGAGCTTCCACGGGACGAGCTGATGCCG 153
QY 61 CTCGAGTCTGGCTTACCGGACGCGCTGGACAAGTACAGCGGAGCACTTGGGCCCGAGAGC 120
Db 154 CTCGAGTCTGGCTTACCGGACGCGCTGGACAAGTACAGCGGAGCACTTGGGCCCGAGAGC 213
QY 121 STKGCTACCTGGAGATCAGCTTGGGCTGCAAGCTTGTCTGGCGACAGGAGGCTTC 180
Db 214 TGGGCTACCTGGAGATCAGCTTGGGCTGCAAGCTTGTCTGGCGACAGGAGGCTTC 273
QY 181 TGCCACCGCAACTGACGCGCGCGCGAGCCCGGAGCGCGGCTTCGCCAGCTAT 240
Db 274 TGCCACCGCAACTGACGCGCGCGCGAGCCCGGAGCGCGGCTTCGCCAGCTAT 333
QY 241 CCCGAGCTGCGCTCTTTCGGGGGCTGCTGCGCGCGGCACTGCTCAAGCGCTGCAAG 300
Db 334 CCCGAGCTGCGCTCTTTCGGGGGCTGCTGCGCGCGGCACTGCTCAAGCGCTGCAAG 393
QY 301 CAGGCGCTGCCAGCTTTCGCGAGTCCGAGCCCGAGCGCGGAGTGTGCGGAGCTTCCAG 360
Db 394 CAGGCGCTGCCAGCTTTCGCGAGTCCGAGCCCGAGCGCGGAGTGTGCGGAGCTTCCAG 453
QY 361 CGCGCGAGCGCTTACAAGTTCCTGCAAGTTCGCTTACTTCAAGGCAATATCTCCCAAA 420
Db 454 CGCGCGAGCGCTTACAAGTTCCTGCAAGTTCGCTTACTTCAAGGCAATATCTCCCAAA 513
QY 421 GCATCGCGCTGCTCAGACCTTCTTCTGAGCATCTGATGACGAATGATGAGAGG 480
Db 514 GCATCGCGCTGCTCAGACCTTCTTCTGAGCATCTGATGACGAATGATGAGAGG 573
QY 481 AACATGCGCATATTAAGAGCTTGTGCTGCGAGGACTTACATTAAGAGCACTTGAAGAC 540
Db 574 AACATGCGCATATTAAGAGCTTGTGCTGCGAGGACTTACATTAAGAGCACTTGAAGAC 632
QY 541 AAGTCATATGAAGCTTGTTCATTCGAGCAGTGTGGGCGCATACAACGCTGAGAACTGAGAG 600
Db 633 AAGTCATATGAAGCTTGTTCATTCGAGCAGTGTGGGCGCATACAACGCTGAGAACTGAGAG 692
QY 601 ACATCCATCACAGACTGAGCTTCCGAGCTTCTTCAAGGCAATATCTTCAAGGCTT 660
Db 693 ACATCCATCACAGACTGAGCTTCCGAGCTTCTTCAAGGCAATATCTTCAAGGCTT 752
QY 661 CTCGAGCTGCGAGGCTTCCAGGAGATCAAGGACTTCAAGGATTTTACCTTTCCATA 720
Db 753 CTCGAGCTGCGAGGCTTCCAGGAGATCAAGGACTTCAAGGATTTTACCTTTCCATA 812

QY 721 GCAGATCATTATGTAGAACTTCTGGNAATGCAAAATACAGTGTGAAGAGAACCTCACCCCA 780
|||
Db 813 GCAG----- 816
QY 781 GTTATAGGAGGCTATCCGGTGTGAGAAATTTGTGCTTACCATATCATTTACTTGCAGTTT 840
Db 817 ----- 816
QY 841 GCCTATTATAAGTTGAACACCTGGAAGANTGAGCCCTGTGCAGTGCAGTATCTGTCT 900
Db 817 -----TGAACACCTGGAAGANTGAGCCCTGTGCAGTGCAGTATCTGTCT 863
QY 901 TTTGATCAGAAATGCAAGGTTCATGACAGAACTCTGTGTATTACAGTACCACAGGAC 960
Db 864 TTTGATCAGAAATGCAAGGTTCATGACAGAACTCTGTGTATTACAGTACCACAGGAC 923
QY 961 ACTTGGGGCTCTCGGATAGCACTTCCAGCCAGACCTGTGATGAGCAGTTTCACTTTTAA 1020
Db 924 ACTTGGGGCTCTCAGATGAGCACTTCCAGCCAGACCTGGAAGCACTTCACTTTTAA 983
QY 1021 GTACACACACTCCAGAGGAGCTGTATGCTTTGCTAAGGAAATATAATGATGATGAT 1080
Db 984 GTACACACACTCCAGAGGAGCTGTATGCTTTGCTAAGGAAATATAATGATGATGAT 1043
QY 1081 GAGGAGAGAGTTTGGAAATATGAGTACCTCTTGGAACTGGAGAGACCAAGC 1134
Db 1044 GAGGAGAGAGTTTGGAAATATGAGTACCTCTTGGAACTGGAGAGACCAAGC 1097

RESULT 3

US-09-962-832-113
; Sequence 113, Application US/09962832
; Patent No. US20020110821A1
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-74
; CURRENT APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR FILING DATE: US/60/235,077
; PRIOR FILING DATE: 2000-09-25
; PRIOR FILING DATE: US/60/235,280
; NUMBER OF SEQ ID NOS: 259
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 113
; LENGTH: 2347
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-832-113

Query Match 40.2%; Score 455.4; DB 10; Length 2347;
Best Local Similarity 65.1%; Pred. No. 2.7e-125;
Matches 723; Conservative 2; Mismatches 368; Indels 18; Gaps 3;
QY 5 CGGCCAATACGACGCTACAGCTTCCGAGCTTCCACCGGAGAGAGAGTGGCGGAGAGCGCTG 64
Db 61 GGGCGCAGTACGAGAGTACAGCTTCCGGGCTTCCCGCCGAGAGCTTATGCCGCTG 120
QY 65 AGTGGCCTACCGGACGCTGGACAGTACAGCGGAGCACTGGGCGGAGAGCGCTG 124
Db 121 CGCGGCTACGCGACGCTCTGGAGCAGTACGAGGAGAGAGTGGCGGAGAGCGC 180
QY 125 GCTACCTGGAGATCAGCTGGGCTGACCCCTTGTGCGGAGAGAGGCTTCTGCC 184
Db 181 GCTACCTGGAGGCGGCTGGGCTGACCCGCTTCTGCGGAGAGAGGCTTCTGCC 240
QY 185 ACCGCAACTCGAGC-----CCGCGCGGAGCGGAGCGGCGGCGGCTCGCCA 235
Db 241 ACGCCAACTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
QY 236 GCTA-----TCCCGAGCTGGCGCTCTTCTGGGGGCGGCTCTGCGGCGGCGGCGGCGCTCA 289

Db 301 ACAGTGGGCTGCGAGCTCGGCTCTTTCGGCGCGCTCTCTGGAGCGAGCGCTGCTGTC 360
QY 290 AGCGCTCAGAGCGGCTGCGAGCTTCCGAGTCCAGTCCAGCCAGCGCGGAGGCTGCTGG 349
Db 361 GGCCTCGAAGCGAGCGCTGCGCGCTTCCAGTGCCTACCGCGCGCGGCGGAGCTGCTGC 420
QY 350 CGGACTTCCAGCGCGCGAGCCCTTACAAAGTTCCTGAGTTCGCTTACTTCAAGGCAATA 409
Db 421 GTGACTTCCAGAGCGGCTGCTTACCACTTACCACTGAGTACCGGCTTCAAGGCTTAACC 480
QY 410 ATCTCCCAAGGCACTGCGCGCTCTCACACCTTCTTACTGAAGCATCTCTGATGACGAAA 469
Db 481 GCGTGGAGAGCGGCTGCGCGCTTACACCTTCTTCCAGAGGAACTCCGAGGCGAGC 540
QY 470 TGAAGAGAGAAACATGGCATATTATAA---GAGCCTTCCCTGCTGCGAGGACTACATTA 526
Db 541 TGACCGCAAGTATCTCAACTTATCAGGGGATGTGGAGCTGCGCGGAGAGTCCCTCA 600
QY 527 AAGACCTGGAAACCAAGTCTATGAAAGCTGTTCATCCGAGCAGTGGGCGGAGTACACG 586
Db 601 CGGACTAGAGGCGGCGCTTACGAGGCGGCTTCTTCCGGGCTGTGAAGCTTCAAGGACT 660
QY 587 GTGAGAACTGGAGAACTCCATCACAGCATGAGAGCTGGCCTTCCCGACTTCTTCAAAG 646
Db 661 CGGGGATTTCCGAGCAGACGCGAGGACATGAGCGGCTTGTGAGAGTACCTGGCAG 720
QY 647 CTTTTAGAGTGTCTGCGAGCTTCCAGGAGTTCAGGAGATCAAGGACTTCAAGGACT 706
Db 721 TCTTTGCGGCTGCTGCGGCTGTGAAGGGGCGGATGAGCAGGCTGAGGACTTCAAGGACT 780
QY 707 TCTACCTTTCCATAGCAGATCATTTATGAGAACTTCTGGAATGCAAAATACAGTGTGAAG 766
Db 781 TCTACCGGCAATAGCAGATCTCTTTCAGAGTCCCTGCACTGCAAGTGGAGTGTGAGG 840
QY 767 AGAAGCTCACCCGAGTATAGAGGCTATCCGGTGTGAAGAAATTTGCTGCTACCATGATC 826
Db 841 CCAATTTGACCCCAATGCGGCTGCTTCTGCGGCAAGTTCGTCGCGGCGGCGGCGGCGG 900
QY 827 ATTACTTGCAGTTTGCCTTATTAAGTTGAACAGCTTGAAGAAATGCAAGCTTGTGCGAG 886
Db 901 ACTACTGCGAGTTTGCCTTACTTAAATGATGATGCGGCGGCGGCTGCGGCGGCGGCGG 960
QY 887 TCAGCTATCTGCTCTTTGATCAGAAATGACAAGTTCATGACAGCAAACTGTGTATTACC 946
Db 961 CCAGCTACATGCTCTTTCAGCCCAAGGACAGCTCATGCAAGCAAACTGTGTATTACC 1020
QY 947 AGTACACAGGAGACACTTGGGCGCTTCCGATGAGCACTTCCAGCCGAGACCTGAGGAG 1006
Db 1021 GGTTCACCGGCTGCTGCGGCTTGAAGAGGAGGAGCTTCCAGCCCGGAGGAGGCGCA 1080
QY 1007 TTCAGTTCTTTAATGTGACCACTCCAGAGGAGCTGTATGACTTTGCTAAGGAAATA 1066
Db 1081 TGCTCTACCAACACAGACCGCGAGCTGCGGAGCTGCTGAGTTCACCCACATGTACC 1140
QY 1067 TAATGATGATGATGAGGAGAAATTTGGA 1097
Db 1141 TGCACTGATGATGATGAGTGGAGTGGAGGA 1171

RESULT 4

US-10-439-388-64
; Sequence 64, Application US/10439388
; Publication No. US20030228617A1
; GENERAL INFORMATION:
; APPLICANT: Aune, Thomas M
; APPLICANT: Olsen, Nancy J
; TITLE OF INVENTION: Method for Predicting Autoimmune Disease
; FILE REFERENCE: 1242/68
; CURRENT APPLICATION NUMBER: US/10/439,388
; CURRENT FILING DATE: 2003-05-16
; PRIOR APPLICATION NUMBER: US 60/381,055
; PRIOR FILING DATE: 2002-05-16

```
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 64
; LENGTH: 2347
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-439-388-64

Query Match      40.2%; Score 455.4; DB 12; Length 2347;
Best Local Similarity 65.1%; Pred. No. 2.7e-125;
Matches 723; Conservative 2; Mismatches 368; Indels 18; Gaps 3;

QY 5 GCGCCCAATACGAGCTACAGCTTCGCGAGCTTCCACGGGACGAGCTGATGCGCTCG 64
Db 61 GGGCGAGTACGAGTACAGCTTCGCGAGCTTCCCGCCGAGGAGCTGATGCGCTCG 120
QY 65 AGTGGCTTCCCGGACGCTGAGCAAGTACAGCGGAGCACTGGGCGGAGAGCTGK 124
Db 121 CCGCGGGGTACGGGACGCTCTGAGCAGTACGAGGAGAGAGCTGGCGGAGCGCG 180
QY 125 GCTACCTGGAGATCAGCTGGGCTGACCGCTTGTGCGGACAGGAGGCTTCTGCG 184
Db 181 GCTACCTGGAGGCGGCTGGGCTGACCGCTTGTGCGGACAGGAGGCTTCTGCG 240
QY 185 ACCGCAACTGCAGCG-----CCGCGCGCAGCGCGGCGCGCGGCGCTCGCA 235
Db 241 ACGCAACTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 300
QY 236 GCTA-----TCCGAGCTGGCTCTTGGGGGCTGCTGCGCGCGCGCTGCTCA 289
Db 301 ACGAGTGGGCTGGAGCTGGGCTCTTGGCGCGCTTCCAGTGCGCTTCCAGTGCG 360
QY 290 AGCGCTGCAACAGCGGCTGCGAGCTTCCGCGAGTCCCGAGCGCGCGAGTCTG 349
Db 361 GCGCTGCAAGCGAGCGCTGCGCGCTTCCAGTGCGCTTCCAGTGCGCTTCCAG 420
QY 350 CGGACTTCCAGCGCGCGGAGCGCTTACAGTCTTGGAGTCTGCTTCAAGGAAATA 409
Db 421 GTGACTTCCAGCGCGGCTGCGCTTCCAGTACCTGCGCTTCCAGTGCGCTTCCAG 480
QY 410 ATCTCCCAAGCATCGCGCTGCTCAGACTTCTTACTGAAGCATCTTGAAGAA 469
Db 481 GCGTGGAGAGCGGCTGGCGCGGCTTACCTTCCAGAGGAAACCGGAGCGAGC 540
QY 470 TGATGAAGAGAAATGCGCATATTATAA---GAGCTGCTGCTGGCGAGACTATTA 526
Db 541 TGACCGCAAGTATCTCAACTACTATCAGGSGATGCTGGAGCTGCGCGAGCTCC 600
QY 527 AAGACTTGGAAACCAAGTCATGAAGCTGTTTCATCCGAGCAGTGGGCGATACA 586
Db 601 CGGACTTGAAGCGCGCGCGCTTACAGGCGCTGTTCTCCGGGCTGTGAAGCT 660
QY 587 GTGAGAACTGGAGAACATCCATCAGACATGAGAGTGGGCTTCCGAGCTTCTCA 646
Db 661 GCGGGAATTCGCGAGCAGCAGGAGGACATGAGGCGGCGCTTGTCAAGTACCT 720
QY 647 CTTTACAGTGTCTGCGAGCTTCCAGGAGTCCAGGAGATCAAGACTTCAAGATT 706
Db 721 TCTTTCGCGGTGCGTGGCGGCTGTGAAGGGGCGGCTGAGCAGGTGAGCTTCAAG 780
QY 707 TCTACTTTCATAGCAGATCATTTATGTAAGTTCGTGAATGCAAAATACAGTGTGA 766
Db 781 TCTACCGCGCCATAGCAGATCTTTTCAGAGTCCCTGAGTCAAGGTGAGCTGT 840
QY 767 AGAAGCTCAGCCAGTTATAGAGCTATCGGTGAGAAATTTGTGCTACCATGATC 826
Db 841 CCAATTTGACCCCAATGCGGCTGCTGCTGAGCAAGTTCGTGGCCACCATGATAC 900
QY 827 ATTACTTCAGTTTCCCTATTATAGTTTGAACGACTTGAAGATGCAAGCTCCCT 886
Db 901 ACTACTGAGTTTGCCTTACTATAGTTTGAATGATGTCGCGAGGCTGCGCGAG 960
QY 887 TCAGTATCTGCTCTTTGATCAGAAATGACAAGGTGATGTCAGCAGAACTGTG 946
```



```

Query Match      13.7%; Score 154.8; DB 14; Length 2600;
Best Local Similarity 53.4%; Pred. No. 2.7e-35;
Matches 436; Conservative 0; Mismatches 347; Indels 33; Gaps 4;

QY 153 CCGCTTGTGCGGACAGCGAGGCTTCTGCGACCGCAACTGCGAGCGCGCGCGAGCC 212
DB 276 CCGCACCAGTGTGCGCGGACTTCCGCTGGAGCTGACCCCGACTGTCCTCCCGAGCCC 335
QY 213 CGAGCCCGCGCGCGCTGCGAGCTATCCGAGCTGCGCTCTTCGGGGCGCTGCTGCG 272
DB 336 GCGCCAGGCTTCCGCGCGCGCGCTGCGAGCTGCGCTCTTCGGGGCGCTTCTGCG 395
QY 273 CCGCGCGAGCTGCTCAAGCGCTCAAGCGCGCTGCGAGCGCTTCCGCGAGTCCCGAGCC 332
DB 396 TCGCGCTGCTTGCCTGCGCGCGCTGCTCGCGCGCGCGCGCGCGCTGCTGCTGCTGCT 446
QY 333 CAGCGCGAGTGTGCGGAGCTTCCGAGCGCGCGCGCTTCCGAGCTGCTGCTGCTGCTGCT 392
DB 447 CAGCGAGAGATGAGCTGAGCTTCCGCAAGCGGAGCGCGCTTCCGAGCTGCTGCTGCTGCT 506
QY 393 TTACTTCAAGCAATAATCTCCCAAGCACTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 452
DB 507 CTACTTCAAGATCAACAGTTGGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 566
QY 453 GCATCTGATGACGAAATGATGAAGAGAAATGAGGATATATAGAGCTGCTGCTGCTGCTGCT 511
DB 567 CAATCTGAGCATGGAATGAGGAAATGAGGAAATGAGGAAATGAGGAAATGAGGAAATGAGG 626
QY 512 --CGAGGACTACATTAAGAGCTGGAACATCCATCAGACATGAGCTGCGGCT 629
DB 687 AGTGGCACTTACTCAGAGGACACCGACAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 746
QY 630 TCCGAGCTTCTCAAGCTTTAGAGTGTCTGCGAGCTGCTGCGAGGTTCCAGGAGAT 689
DB 747 GCAAGATACATTTGTGGCTATGAGGAGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 806
QY 690 CAGGAGCTTCAA-----GGATTCTACCTTTCCATAGCAGATCATTA 731
DB 807 CGATGGCTACAACTACCTTGAAGTACAGCTGAGCTTCCAGGCTTCCAGGCTTCCAGGCTT 866
QY 732 TGTAGAAGTCTGGAATCAAAATACAGTGTGAAGAGAACCT---CACCCAGTTATAGG 788
DB 867 CATCCAGGTCCTCAACTGTAGCAGAACTGTGTACGAGGCTGCTTCCACCCAGGCTG 926
QY 789 AGGCTATCCGCTTGAAGAAATTTGTGGTACCAATGATCATTTACTTGCAGTTTGCCTATTA 848
DB 927 AGAGAAGCCCTTTGAGAGCTTCTCCATCGCATTAATATATCTGAGTTTGCCTACTA 986
QY 849 TAAATTGAACGACTGGAAGAAATGACGCGCTGCGAGCTGAGTATCTGCTCTTTGATCA 908
DB 987 TAACATTTGGAAATTTATACACAGCTGCTGGAATGCTGCAAGAGCTTCTCTCTTCTCC 1046
QY 1047 CAATGACGAGGTGATGAACCAAAATTTGGCCTATTA 1082

```

```

RESULT 8
US-10-045-815-1
; Sequence 1, Application US/10045815
; Publication No. US20020160498A1
; GENERAL INFORMATION:
; APPLICANT: Wachwa, Renu
; APPLICANT: Sugihara, Takashi
; APPLICANT: Ohide, Akiko
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; FILE REFERENCE: 06501-091001
; CURRENT APPLICATION NUMBER: US/10/045,815
; CURRENT FILING DATE: 2001-10-26

```

```

; PRIOR APPLICATION NUMBER: PCT/JP00/02731
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: JP 11/118806
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2829
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (52)...(1140)
; US-10-045-815-1

```

```

Query Match      13.7%; Score 154.8; DB 14; Length 2829;
Best Local Similarity 53.4%; Pred. No. 2.8e-35;
Matches 436; Conservative 0; Mismatches 347; Indels 33; Gaps 4;

QY 153 CCGCTTGTGCGGACAGCGAGGCTTCTGCGACCGCAACTGCGAGCGCGCGCGAGCC 212
DB 276 CCGCACCAGTGTGCGCGGACTTCCGCTGGAGCTGACCCCGACTGTCCTCCCGAGCCC 335
QY 213 CGAGCCCGCGCGCGCTGCGAGCTATCCGAGCTGCGCTCTTCGGGGCGCTGCTGCG 272
DB 336 GCGCCAGGCTTCCGCGCGCGCGCTGCGAGCTGCGCTCTTCGGGGCGCTTCTGCG 395
QY 273 CCGCGCGAGCTGCTCAAGCGCTGCAAGCGCGCTGCGAGCGCTTCCGCGAGTCCCGAGCC 332
DB 396 TCGCGCTGCTTGCCTGCGCGCGCTGCTCGCGCGCGCGCGCGCTGCTGCTGCTGCT 446
QY 333 CAGCGCGAGTGTGCGGAGCTTCCGAGCGCGCGCGCTTCCGAGCTGCTGCTGCTGCTGCT 392
DB 447 CAGCGAGAGATGAGCTGAGCTTCCGCAAGCGGAGCGCGCTTCCGAGCTGCTGCTGCTGCT 506
QY 393 TTACTTCAAGCAATAATCTCCCAAGCACTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 452
DB 507 CTACTTCAAGATCAACAGTTGGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 566
QY 453 GCATCTGATGACGAAATGATGAAGAGAAATGAGGATATATAGAGCTGCTGCTGCTGCTGCT 511
DB 567 CAATCTGAGCATGGAATGAGGAAATGAGGAAATGAGGAAATGAGGAAATGAGGAAATGAGG 626
QY 512 --CGAGGACTACATTAAGAGCTGGAACATCCATCAGACATGAGCTTTCATCCGAGC 569
DB 627 GAAGGAGCGGACTTCCAGGATCTTGAAGTCAACCCATATGCAAGAAATTTTCGAGTGG 686
QY 570 AGTGGCGGATACAAAGCTGAGAACTGGAAGAAATGAGGAAATGAGGAAATGAGGAAATGAGG 629
DB 687 AGTGGCACTTACTCAGAGGAAACAGCCACAGGAAAGCTGTGCGCGGCTGAGAGGCGGCT 746
QY 630 TCCGAGCTTCTCAAGCGCTTTAGAGTGTCTGCGAGCTGCTGCGAGGTTCCAGGAGAT 689
DB 747 GCAAGAAATCTTGTGGCTATGAGGAGTGGCGTGGCTGCTGCGAAGGCGCTTATGACTA 806
QY 690 CAGGAGCTTCAA-----GGATTCTACCTTTCCATAGCAGATCATTA 731
DB 807 CGATGGCTACAACTACCTTGAAGTACAGCTGAGCTTCCAGGCTTCCAGGCTTCCAGGCTT 866
QY 732 TGTAGAAGTCTGGAATGCAAAATACAGTGTGAAGAGAACCT---CACCCAGTTATAGG 788
DB 867 CATCCAGGTCCTCAACTGTAGCAGAACTGTGTACGAGGCTGCTTCCACCCAGGCTG 926
QY 789 AGGCTATCCGCTTGAAGAAATTTGTGGTACCAATGATCATTTACTTGCAGTTTGCCTATTA 848
DB 927 AGAGAAGCCCTTTGAGAGCTTCTCCATCGCATTAATATATCTGAGTTTGCCTACTA 986
QY 849 TAAATTGAACGACTGGAAGAAATGACGCGCTGCGAGCTGAGTATCTGCTCTTTGATCA 908
DB 987 TAACATTTGGAAATTTATACACAGCTGCTGGAATGCTGCAAGAGCTTCTCTCTTCTCC 1046
QY 909 GAATGACAGGTGATGACCAAAATTTGGCCTATTA 944

```

Db 1047 CAATGACGAGTGTGATGAACCAAAATTTGGCTATTA 1082

RESULT 9

US-10-027-632-161474/c
; Sequence 161474, Application US/10027632
; Publication No. US20030204075A9

GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

; FILE OF INVENTION: Polymorphisms in the Human Genome

; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027,632

; CURRENT FILING DATE: 2002-04-30

; PRIOR APPLICATION NUMBER: US 60/218,006

; PRIOR FILING DATE: 2000-07-12

; PRIOR APPLICATION NUMBER: US 60/198,676

; PRIOR FILING DATE: 2000-04-20

; PRIOR APPLICATION NUMBER: US 60/193,483

; PRIOR FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: US 60/185,218

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/167,363

; PRIOR FILING DATE: 1999-11-23

; PRIOR APPLICATION NUMBER: US 60/156,358

; PRIOR FILING DATE: 1999-09-28

; PRIOR APPLICATION NUMBER: US 60/146,002

; PRIOR FILING DATE: 1999-08-09

; NUMBER OF SEQ ID NOS: 325720

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 161474

; LENGTH: 892

; TYPE: DNA

; ORGANISM: Human

US-10-027-632-161474

Query Match

Best Local Similarity 13.6%; Score 153.8; DB 13; Length 892;

Matches 155; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 394 TACTTCAAGGCAATAATCTCCCAAGCCATCGCGCTGCTCAGACCTTTCTACTGAAG 453

Db 228 TCCCTTTAGGCAATAATCTCCCAAGCCATCGCGCTGCTCAGACCTTTCTACTGAAG 169

QY 454 CATCCTGATGACGAATGATGAGGAAACATGCGCATATTAAGAGCCTGCTGTGTC 513

Db 168 CATCCTGATGAGAAATGATGAGGAAACATGCGCATATTAAGAGCCTGCTGTGTC 109

QY 514 GAGGACTACATTAAGACCTGGAACCAAGTCATATGAA 552

Db 108 GAGGACTACATTAAGACCTGGAACCAAGTCATATGAA 70

RESULT 10

US-10-027-632-161475/c

; Sequence 161475, Application US/10027632

; Publication No. US20030204075A9

GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

; FILE OF INVENTION: Polymorphisms in the Human Genome

; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027,632

; CURRENT FILING DATE: 2002-04-30

; PRIOR APPLICATION NUMBER: US 60/218,006

; PRIOR FILING DATE: 2000-07-12

; PRIOR APPLICATION NUMBER: US 60/198,676

; PRIOR FILING DATE: 2000-04-20

; PRIOR APPLICATION NUMBER: US 60/193,483

; PRIOR FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: US 60/185,218

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/167,363

; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 161475

; LENGTH: 892

; TYPE: DNA

; ORGANISM: Human

US-10-027-632-161475

Query Match

Best Local Similarity 13.6%; Score 153.8; DB 13; Length 892;

Matches 155; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 394 TACTTCAAGGCAATAATCTCCCAAGCCATCGCGCTGCTCAGACCTTTCTACTGAAG 453

Db 228 TCCCTTTAGGCAATAATCTCCCAAGCCATCGCGCTGCTCAGACCTTTCTACTGAAG 169

QY 454 CATCCTGATGACGAATGATGAGGAAACATGCGCATATTAAGAGCCTGCTGTGTC 513

Db 168 CATCCTGATGAGAAATGATGAGGAAACATGCGCATATTAAGAGCCTGCTGTGTC 109

QY 514 GAGGACTACATTAAGACCTGGAACCAAGTCATATGAA 552

Db 108 GAGGACTACATTAAGACCTGGAACCAAGTCATATGAA 70

RESULT 11

US-10-027-632-161476/c

; Sequence 161476, Application US/10027632

; Publication No. US20030204075A9

GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

; FILE OF INVENTION: Polymorphisms in the Human Genome

; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027,632

; CURRENT FILING DATE: 2002-04-30

; PRIOR APPLICATION NUMBER: US 60/218,006

; PRIOR FILING DATE: 2000-07-12

; PRIOR APPLICATION NUMBER: US 60/198,676

; PRIOR FILING DATE: 2000-04-20

; PRIOR APPLICATION NUMBER: US 60/193,483

; PRIOR FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: US 60/185,218

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/167,363

; PRIOR FILING DATE: 1999-11-23

; PRIOR APPLICATION NUMBER: US 60/156,358

; PRIOR FILING DATE: 1999-09-28

; PRIOR APPLICATION NUMBER: US 60/146,002

; PRIOR FILING DATE: 1999-08-09

; NUMBER OF SEQ ID NOS: 325720

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 161476

; LENGTH: 892

; TYPE: DNA

; ORGANISM: Human

US-10-027-632-161476

Query Match

Best Local Similarity 13.6%; Score 153.8; DB 13; Length 892;

Matches 155; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 394 TACTTCAAGGCAATAATCTCCCAAGCCATCGCGCTGCTCAGACCTTTCTACTGAAG 453

Db 228 TCCCTTTAGGCAATAATCTCCCAAGCCATCGCGCTGCTCAGACCTTTCTACTGAAG 169

QY 454 CATCCTGATGACGAATGATGAGGAAACATGCGCATATTAAGAGCCTGCTGTGTC 513

Db 168 CATCTGATGAYGAAATGATGAAGAGGAACATGGCATATATATAAGAGCCTGCTGTGTC 109
Qy 514 GAGGACTACATTAAAGACCTGGAAACCAAGTCATATGAA 552
Db 108 GAGGACTACATTAAAGACCTGGAAACCAAGTCATATGAA 70

RESULT 12
US-10-027-632-161474/c
; Sequence 161474, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 161474
; LENGTH: 892
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-161474

Query Match 13.6%; Score 153.8; DB 14; Length 892;
Best Local Similarity 97.5%; Pred. No. 3e-35;
Matches 155; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Qy 394 TACTTCAAGGCAAAATAATCTCCCAAGCCATCCCGCTGCTCAGCTTCTACTGAAG 453
Db 228 TCCCTTTAGGCAAAATAATCTCCCAAGCCATCCCGCTGCTCAGCTTCTACTGAAG 169
Qy 454 CATCTGATGACGAAATGATGAAGAGGAACATGGCATATATATAAGAGCCTGCTGTGTC 513
Db 168 CATCTGATGAYGAAATGATGAAGAGGAACATGGCATATATATAAGAGCCTGCTGTGTC 109
Qy 514 GAGGACTACATTAAAGACCTGGAAACCAAGTCATATGAA 552
Db 108 GAGGACTACATTAAAGACCTGGAAACCAAGTCATATGAA 70

RESULT 13
US-10-027-632-161475/c
; Sequence 161475, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218

; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 161475
; LENGTH: 892
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-161475
Query Match 13.6%; Score 153.8; DB 14; Length 892;
Best Local Similarity 97.5%; Pred. No. 3e-35;
Matches 155; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Qy 394 TACTTCAAGGCAAAATAATCTCCCAAGCCATCCCGCTGCTCAGCTTCTACTGAAG 453
Db 228 TCCCTTTAGGCAAAATAATCTCCCAAGCCATCCCGCTGCTCAGCTTCTACTGAAG 169
Qy 454 CATCTGATGACGAAATGATGAAGAGGAACATGGCATATATATAAGAGCCTGCTGTGTC 513
Db 168 CATCTGATGAYGAAATGATGAAGAGGAACATGGCATATATATAAGAGCCTGCTGTGTC 109
Qy 514 GAGGACTACATTAAAGACCTGGAAACCAAGTCATATGAA 552
Db 108 GAGGACTACATTAAAGACCTGGAAACCAAGTCATATGAA 70

RESULT 14
US-10-027-632-161476/c
; Sequence 161476, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 161476
; LENGTH: 892
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-161476

Query Match 13.6%; Score 153.8; DB 14; Length 892;
Best Local Similarity 97.5%; Pred. No. 3e-35;
Matches 155; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Qy 394 TACTTCAAGGCAAAATAATCTCCCAAGCCATCCCGCTGCTCAGCTTCTACTGAAG 453
Db 228 TCCCTTTAGGCAAAATAATCTCCCAAGCCATCCCGCTGCTCAGCTTCTACTGAAG 169
Qy 454 CATCTGATGACGAAATGATGAAGAGGAACATGGCATATATATAAGAGCCTGCTGTGTC 513

```

168 CATCTGATGAYGAAATGATGAAGAGAACATGGCATATTATAAGAGCCTGCGTGGCC 109
514 GAGGACTACATTAAGACCTGGAAACCAAGTCATATGAA 552
108 GAGGACTACATTAAGACCTGGAAACCAAGTCATATGAA 70

RESULT 15
US-10-045-815-7
; Sequence 7, Application US/10045815
; Publication No. US20020160498A1
; GENERAL INFORMATION:
; APPLICANT: Wadhwah, Renu
; APPLICANT: Sugihara, Takashi
; APPLICANT: Ohide, Akiko
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; FILE REFERENCE: 06501-091001
; CURRENT APPLICATION NUMBER: US/10/045,815
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: PCT/JP00/02731
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: JP 11/118806
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 2322
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (12)...(1637)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 2282
; OTHER INFORMATION: n = A,T,C or G
US-10-045-815-7

Query Match 13.0%; Score 147.4; DB 14; Length 2322;
Best Local Similarity 54.6%; Pred. No. 4,1e-33;
Matches 421; Conservative 0; Mismatches 316; Indels 34; Gaps 5;

Qy 200 CCGCGCGCGCAGCCGAGCC-CGCGCGCGCGCTCGCAGCTATCCGAGCTGCGCCTCTTC 258
Db 288 CCGAGACCCAGCTGAGCCAGGACCCGCGCGCGCGCTCGCAGCTATCCGAGCTGCGCCTCTTC 347
Qy 259 GGGGGCTGTGCGCGCGCGCAGCTGCGCTCAAGCGCTGCAAGCAGGCGCTTSCCAGCGCTTC 318
Db 348 GGAGCGGTGTGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCGCGCGCTGCGCGCAC 407
Qy 319 CGCCAGTCCAGCCAGCCGCGGAGGTGTGCGGAGCTTCCAGCGCGCGGAGCCCTACAAG 378
Db 408 TTGCTG-----AGTGAAGAACTGGAGCTGAGTTCACAGCGGAGCCCGGTACAAC 458
Qy 379 TTCTGTGAGTTCGGTCTTCAAGGCAATAATCTCCCCAAAGCCATCGCGCTGTCTCAC 438
Db 459 TACCTGAGGTGCGCTATTTCAGATTAACAGCTGGAGAGGTGTGCTGCGGCACAC 518
Qy 439 ACCTTTCTACTGAAGCATCTGTATGAGAAATGATGAAGAGGAAACATGGCATATTATAAG 498
Db 519 ACCTTCTTTTGGGCAATCTCTGAGCAGATGGAGATGGGAGAACCTCGACTATTACCAA 578
Qy 499 AGCTGTGCTGGTG--CCGAGGACTACATTAAAGACCTGGAAACCAAGTCATATGAAAGC 555
Db 579 ACCATGCTGGGTGAGAGAGGAGCAGACTTCAGGGATCTCGAGGCCAAGCCCATATGCAT 638
Qy 556 CTGTTTCATCCAGCAGTGGCGGCATACAAACGGTGAGAACTGGAGAACATCCATCACAGAC 615
Db 639 GAGTTTCGGCTGGGGTACGACTCTACTCAGAGGAGAAAGCCACAGGAAGCTGTGCCCCAC 698
Qy 616 ATGAGCTGGCGCTTCCGAGCTTTTCAAGACCTTTTACGAGTGTCTCGCAGCGCTGCGAG 675

```

```

699 CTGGAGGGCGGCACTGCAAGAGTACTTTCTGGCCGATGAGGAGTGCCTGCGCTCTGCAG 758
Qy 676 GGTTCAGGGAGATCAAGGACTTCAA-----GGATTTCTACCTTTCC 717
Db 759 GGGCCCTATGACTAGCAGCGGTACAACTACCTAGACTACAGCGGTGACTCTTTCAGGCC 818
Qy 718 ATAGCAGATCATTTATGTAGAAGTTCTTGAAGTGTGGAATGCAAAATACAGTGTGAAGAACT---C 774
Db 819 ATCAGAGATCATTTACGTCAGGTCCAGTCCCTCACTGTAAGCAGAACTGTGTACGAGCTGGCT 878
Qy 775 ACCCCAGTTATAGGAGGCTATCCGTTTCAGAAATTTGGGTACCATGTATCATTTACTTG 834
Db 879 TCCACCCAGTAGGAGAAAGCCCTTTGAAGACTTTCCTCCCTTTCACACTATAATTACCTA 938
Qy 835 CAGTTTGGCTATTATAAGTGTGAACGACCTGAAGAAATGCAAGCCCTGTGCGAGTCAGCTAT 894
Db 939 CAGTTTGGCTACTACACACATTTGGGAACCTATACAGAGCTATTGAATGTGCCAAGACCTAC 998
Qy 895 CTGCTCTTTGATCAGAAATGACAAGTCTATGAGCAGAGAACCTGGTGTATTATAC 945
Db 999 CTCCTCTTTCTTCCCAATGATGAGGTGATGCACCAAGATCTGGCTTATTAC 1049

```

Search completed: January 30, 2004, 19:07:44
Job time : 395.871 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2004, 08:05:34 ; Search time 63.9336 Seconds
(without alignments)
7828.876 Million cell updates/sec

Title: US-09-729-674-1_COPY_132_1265

Perfect score: 1134

Sequence: 1 gggcgcccaatacgaacg.....tgggaactggaggagaccagc 1134

Scoring table: IDENTITY NUC

Gapop 10_0, Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/prodata/1/ina/PCUS_COMB.seq: *
6: /cgn2_6/prodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	126	11.1	618	3	US-09-328-111-215 Sequence 215, Appl
C 2	54.4	4.8	30001	1	US-08-125-468-1 Sequence 1, Appl
C 3	54.4	4.8	30001	2	US-08-474-933-1 Sequence 1, Appl
4	52	4.6	3624	1	US-07-951-715A-6 Sequence 6, Appl
5	52	4.6	3624	2	US-08-459-448A-6 Sequence 6, Appl
6	52	4.6	3624	3	US-08-459-595A-6 Sequence 6, Appl
7	52	4.6	3624	3	US-08-459-504B-6 Sequence 6, Appl
8	52	4.6	3624	3	US-08-459-444-6 Sequence 6, Appl
9	52	4.6	3624	3	US-09-053-549-7 Sequence 7, Appl
10	52	4.6	3624	4	US-09-547-422-6 Sequence 3976, Ap
11	50.6	4.5	942	4	US-09-252-991A-3976 Sequence 4046, Ap
12	49.8	4.4	13542	3	US-09-105-537-30 Sequence 30, Appl
13	49.8	4.4	36778	3	US-09-105-537-5 Sequence 5, Appl
14	49.8	4.4	38506	3	US-09-320-878-19 Sequence 19, Appl
15	49.8	4.4	38506	4	US-09-141-908-1 Sequence 1, Appl
16	49.8	4.4	38506	4	US-09-657-440-19 Sequence 19, Appl
17	49.8	4.4	38506	4	US-08-324-243-35 Sequence 35, Appl
18	49.2	4.3	2481	1	US-08-532-390-35 Sequence 35, Appl
19	49.2	4.3	2481	2	US-08-717-294-35 Sequence 35, Appl
20	49.2	4.3	2481	3	US-08-717-294-35 Sequence 35, Appl
21	49.2	4.3	2481	5	PCI-US95-11511-35 Sequence 35, Appl
22	49.2	4.3	2571	4	US-09-552-950-4 Sequence 4, Appl
23	49	4.3	5228	4	US-09-428-711A-15 Sequence 15, Appl
24	48.4	4.3	1162	2	US-08-726-306A-52 Sequence 52, Appl
25	48.2	4.3	414	4	US-09-252-991A-2183 Sequence 2183, Ap
26	48.2	4.3	1635	4	US-09-252-991A-2493 Sequence 2493, Ap
27	48.2	4.3	1365	4	US-09-252-991A-9230 Sequence 9230, Ap

28 48.2 4.3 2805 4 US-09-252-991A-9064 Sequence 9064, Ap
 29 48 4.2 567 4 US-09-252-991A-13282 Sequence 13282, A
 30 48 4.2 1425 4 US-09-252-991A-12820 Sequence 12820, A
 31 48 4.2 1428 4 US-09-252-991A-13677 Sequence 12677, A
 32 48 4.2 1611 4 US-09-252-991A-13122 Sequence 13122, A
 33 48 4.2 4403765 3 US-09-103-840A-2 Sequence 2, Appl
 34 48 4.2 4411529 3 US-09-103-840A-1 Sequence 1, Appl
 35 47.8 4.2 1155 4 US-09-252-991A-1572 Sequence 1572, Ap
 36 47.8 4.2 4239 4 US-09-252-991A-1619 Sequence 1619, Ap
 37 47.8 4.2 4353 4 US-09-252-991A-1547 Sequence 1547, Ap
 38 47.5 4.2 1288 1 US-08-440-856A-9 Sequence 9, Appl
 39 47.2 4.2 1637 4 US-09-252-991A-15493 Sequence 15493, A
 40 47.2 4.2 2157 4 US-09-252-991A-13294 Sequence 15294, A
 41 47.2 4.2 2361 4 US-09-252-991A-13546 Sequence 15546, A
 42 47 4.1 891 4 US-09-252-991A-1176 Sequence 1176, App
 43 47 4.1 987 4 US-09-252-991A-14271 Sequence 14271, A
 44 47 4.1 996 2 US-08-924-759-21 Sequence 21, Appl
 45 47 4.1 996 3 US-09-248-335-21 Sequence 21, Appl

ALIGNMENTS

RESULT 1
 US-09-328-111-215/c
 ; Sequence 215, Application US/09328111
 ; Patent No. 6262333
 ; GENERAL INFORMATION:
 ; APPLICANT: Endege, Wilson O.
 ; APPLICANT: Steinmann, Kathleen E.
 ; APPLICANT: Astle, Jon H.
 ; APPLICANT: Burgess, Christopher C.
 ; APPLICANT: Bushnell, Steven E.
 ; APPLICANT: Carroll III, Eddie
 ; APPLICANT: Catino, Theodore J.
 ; APPLICANT: Derti, Adnan
 ; APPLICANT: Ford, Donna M.
 ; APPLICANT: Lewis, Marcia E.
 ; APPLICANT: Monahan, John E.
 ; APPLICANT: Schlegel, Robert
 ; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
 ; TITLE OF INVENTION: PRODUCTS
 ; FILE REFERENCE: CCD-257 (US)
 ; CURRENT APPLICATION NUMBER: US/09/328,111
 ; NUMBER OF SEQ ID NOS: 850
 ; EARLIER FILING DATE: 1999-06-08
 ; EARLIER APPLICATION NUMBER: US 60/088,801
 ; NUMBER OF SEQ ID NOS: 850
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 215
 ; LENGTH: 618
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(618)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-328-111-215

Query Match 11.1%; Score 126; DB 3; Length 618;
 Best Local Similarity 85.9%; Pred. No. 1.4e-21;
 Matches 165; Conservative 0; Mismatches 24; Indels 3; Gaps 3;
 QY 945 CCAGTACCACAGGACACTTGGGGCCCTCTCGGATGAGCACTTCCAGCCCAAGCCTGAAGC 1004
 534 CNAAGTCCCNAGNCCANTGGGCGCTTNGAATGAGCANTTCCAGCCCAAGCTTAAANC 475
 QY 1005 AGTTCACTTCTTAAATGTGACCA-CAGTCCAGAGAGGC-TGTATGACTTT-CCTAAGGA 1061
 474 AGTTCACTTCTTAAATGTGNCACCACTCCAGANGAAGCTTGTATGACTTTGNTAAGGA 415
 QY 1062 AAATATATGATGATGATGAGGAGAGTGTGGGAATATGTGATGACCTTTGGAAC 1121

Db 6020 GACCTCCCGGAGGACACCTGCTGACACGAGCCCGGACCCCTGCTGCTGCTC 5961
QY 280 CACTGCTCAAGCGTTCAGAGGCGCTGCGAGCTTCGCGGCACTCCAGCCCGGCGC 339
Db 5960 CGGTTCCCGGAGCCCGGCGGAGTGGGCGTGTGGACACGCTGGACACCGA---CGGCGCC 5904
QY 340 GAGGTGCTGGCGACTTCAGAGCGCGGAGCCCTACAGTTCTTCAGTTGCT 393
Db 5903 GAGGACCGGAGTGTGCTCCGCGCGCCCTGCGCGACAGATCGGCGCGGGCGCT 5850

RESULT 4

US-07-951-715A-6
; Sequence 6, Application US/07951715A
; Patent No. 5825136
; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
; APPLICANT: Desai, Nalini M.
; APPLICANT: Lewis, Kelly S.
; APPLICANT: Kramer, Vance C.
; APPLICANT: Warren, Gregory W.
; APPLICANT: Evola, Stephen V.
; APPLICANT: Crossland, Lyle D.
; APPLICANT: Wright, Martha S.
; APPLICANT: Merlin, Ellis J.
; APPLICANT: Launis, Karen L.
; APPLICANT: Rothstein, Steven J.
; APPLICANT: Bowman, Cindy G.
; APPLICANT: Dawson, John L.
; APPLICANT: Dunder, Erik M.
; APPLICANT: Pace, Gary M.
; APPLICANT: Suttie, Janet L.
; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
; TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: New York
; COUNTRY: USA
; ZIP: 10532

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30B
CURRENT APPLICATION DATA: US/07/951,715A
FILING DATE: 25-SEP-1992
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/772,027
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Sprull, W. Murray
REGISTRATION NUMBER: 32,943
REFERENCE/DOCKET NUMBER: S-18805/A/CGC 1577/CIP
TELEPHONE: (919)541-8615
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 3624 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Synthetic DNA"
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: CDS

LOCATION: 1..3621
OTHER INFORMATION: /product= "Full-length, maize
OTHER INFORMATION: optimized cry1B"
OTHER INFORMATION: /note= "Disclosed in Figure 6."
US-07-951-715A-6
Query Match 4.6%; Score 52; DB 1; Length 3624;
Best Local Similarity 47.5%; Pred. No. 0.0027;
Matches 151; Conservative 1; Mismatches 166; Indels 0; Gaps 0;
QY 100 GGGGAGCACTGGGCGGAGAGCTKGGCTACTCTGGAGATCAGCTGGGCTGCGACCGCTTG 159
Db 181 GGCAGCTGTGGCCCCCGCGCGGACCACTGAGGATCTTCTGGAGCAGCTGGAGCAG 240
QY 160 CTGCGGACAGCGAGGCTTTTCCACCGCAATGAGAGCGCGCGCGCGAGCCGAGCCG 219
Db 241 CTGATCAACAGCAGATCAGCGAGAACGCCCGCAACACCGCCCTGCGCGCGG 300
QY 220 GCGCGCGCTCGCCAGCTATCCGAGCTGGCTCTTCGGGGCGCTGCTGCGCGCGCG 279
Db 301 CTGGCGAGACAGTTCCCGCGCTACCCAGAGCTGGAGACTGGCTGGAGAACCGGAC 360
QY 280 CACTGCTCAAGCGCTGCAAGAGCGGCTCCAGAGCTTCGCGAGTTCGCGAGCCGAGCCG 339
Db 361 GACCGCGGACCGCGAGGCTGTGTACACCCAGTACATCGCCCTGGAGCTGGACTTCTG 420
QY 340 GAGGTGCTGGGCGACTTCCAGCGCGCGAGCCCTACAGTTCTTCGAGTTGCTTACTTC 399
Db 421 AACGCGATGCGCCCTGTTGCGCATCCGCAACAGAGTGCCCTGCTGATGTTGCTACGCC 480
QY 400 AAGGCAATAATCTCCCC 417
Db 481 CAGCGCGCAACTGCGAC 498

RESULT 5

US-08-459-448A-6
; Sequence 6, Application US/08459448A
; Patent No. 5959336
; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
; APPLICANT: Desai, Nalini M.
; APPLICANT: Lewis, Kelly S.
; APPLICANT: Kramer, Vance C.
; APPLICANT: Warren, Gregory W.
; APPLICANT: Evola, Stephen V.
; APPLICANT: Crossland, Lyle D.
; APPLICANT: Wright, Martha S.
; APPLICANT: Merlin, Ellis J.
; APPLICANT: Launis, Karen L.
; APPLICANT: Rothstein, Steven J.
; APPLICANT: Bowman, Cindy G.
; APPLICANT: Dawson, John L.
; APPLICANT: Dunder, Erik M.
; APPLICANT: Pace, Gary M.
; APPLICANT: Suttie, Janet L.
; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
; TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5859336artis Corporation
; STREET: Patent & Trademark Dept., 520 White Plains
; STREET: Rd., POB 2005
; CITY: Tarrytown
; STATE: New York
; COUNTRY: USA
; ZIP: 10591-9005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:


```

; APPLICATION NUMBER: US/08/459,448A
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/951,715
; FILING DATE: 25-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Gary M.
; REGISTRATION NUMBER: 40403
; REFERENCE/DOCKET NUMBER: CGC 1577/CIP/DIV4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8582
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3624 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Synthetic DNA"
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3621
; OTHER INFORMATION: /product= "Full-length, maize
; OTHER INFORMATION: optimized cryiB"
; OTHER INFORMATION: /note= "Disclosed in Figure 6."
;
US-08-459-448A-6

Query Match 4.6%; Score 52; DB 2; Length 3624;
Best Local Similarity 47.5%; Pred. No. 0.0027;
Matches 151; Conservative 1; Mismatches 166; Indels 0; Gaps 0;

QY 100 GCGGAGCTGGCGGCGAGCGTGGCTACCTGGAGATCAGCTCGGCTGCACGGCTTG 159
Dd 181 GCGGAGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
QY 160 CTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 219
Dd 241 CTGATCAACGAGATACACGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
QY 220 GCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 279
Dd 301 CTGGGCGGAGCTTCCGCGCTACCGAGAGCGGCTGGAGGAGTGGCTGGAGAACGCGGAC 360
QY 280 CACTGCTCAAGGCTGCGAGGCGGCTGGAGGCGGCTGGAGGCGGCTGGAGGCGG 339
Dd 361 GACGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
QY 340 GAGGTGCTGGCGGAGCTTCCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 399
Dd 421 AAGCGCATGCGGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
QY 400 AAGGCAAAATATCTCCCG 417
Dd 481 CAGGCGCGCAACCTGCAC 498

RESULT 6
US-08-459-595A-6
; Sequence 6, Application US/08459595A
; Patent No. 6018104
; GENERAL INFORMATION:
; APPLICANT: Michael G.
; APPLICANT: Desai, Nalini M.
; APPLICANT: Lewis, Kelly S.
; APPLICANT: Kramet, Vance C.
; APPLICANT: Warren, Gregory W.
; APPLICANT: Evola, Stephen V.
;

```

```

; APPLICANT: Crossland, Lyle D.
; APPLICANT: Wright, Martha S.
; APPLICANT: Merlin, Ellis J.
; APPLICANT: Launis, Karen L.
; APPLICANT: Rothstein, Steven J.
; APPLICANT: Bowman, Cindy G.
; APPLICANT: Dawson, John L.
; APPLICANT: Dunder, Erik M.
; APPLICANT: Pace, Gary M.
; APPLICANT: Suttie, Janet L.
; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
; TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6018104artis Corporation
; STREET: Patent & Trademark Dept., 520 White Plains
; STREET: Rd., POB 2005
; CITY: Tarrytown
; STATE: New York
; COUNTRY: USA
; ZIP: 10591-9005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/459,595A
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/951,715
; FILING DATE: 25-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Gary M.
; REGISTRATION NUMBER: 40403
; REFERENCE/DOCKET NUMBER: CGC 1577/CIP/DIV3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8582
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3624 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Synthetic DNA"
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3621
; OTHER INFORMATION: /product= "Full-length, maize
; OTHER INFORMATION: optimized cryiB"
; OTHER INFORMATION: /note= "Disclosed in Figure 6."
;
US-08-459-595A-6

Query Match 4.6%; Score 52; DB 3; Length 3624;
Best Local Similarity 47.5%; Pred. No. 0.0027;
Matches 151; Conservative 1; Mismatches 166; Indels 0; Gaps 0;

QY 100 GCGGAGCTGGCGGCGAGAGCGTGGCTACCTGGAGATCAGCTCGGCTGCACCGCTTG 159
Dd 181 GCGGAGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
QY 160 CTGCGCGGAGCGGAGGCGCTTCTGCCACCGCAATGTGAGCGGCGGCGGCGGCGGCGG 219
Dd 241 CTGATCAACGAGAGATCACCGAGAACGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
QY 220 GCGCGCGGCTGGCGGAGCTTCCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 279

```

Db 301 CTGGGCGACAGCTTCCGGCGCTACACGACAGAGCTGGAGGACTGGCTGGAGAACCGGAC 360
QY 280 CACTGCCTCAAGCGCTGAACAGAGGCTGCAGAGCTTCCGCACTCCAGCCAGCGGC 339
Db 361 GACCCCGCACCAGCGTCTGTACACCAAGTACATGCGCTGGAGCTGGACTTCTG 420
QY 340 GAGTGTCTGGCGGACTTCCAGCGCGGAGAGCTTCAAGTTCTCTGCAAGTTCTTACTTC 399
Db 421 AACCCATGCCCTTTCGCCATCCGCAACAGAGGTGCCCTGCTGATGGTGTACGCC 480
QY 400 AAGGCAATAATCTCCCC 417
Db 481 CAGCGCGCAACCTGCAC 498

RESULT 7
US-08-459-504B-6
; Sequence 6, Application US/08459504B
; Patent No. 6075185
; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
; APPLICANT: Desai, Nalini M.
; APPLICANT: Lewis, Kelly S.
; APPLICANT: Kramer, Vance C.
; APPLICANT: Warren, Gregory W.
; APPLICANT: Evola, Stephen V.
; APPLICANT: Crossland, Lyle D.
; APPLICANT: Wright, Martha S.
; APPLICANT: Merlin, Ellis J.
; APPLICANT: Launis, Karen L.
; APPLICANT: Rothstein, Steven J.
; APPLICANT: Bowman, Cindy G.
; APPLICANT: Dawson, John L.
; APPLICANT: Dunder, Erik M.
; APPLICANT: Pace, Gary M.
; APPLICANT: Suttle, Janet L.
; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
; TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 6075185artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,504B
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/459,595
; FILING DATE: 02-JUN-1995
; APPLICATION NUMBER: US 07/951,715
; FILING DATE: 25-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC1577/CIP/DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8587
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 3624 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Synthetic DNA"
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3621
; OTHER INFORMATION: /product= "Full-length, maize
; OTHER INFORMATION: Optimized cryiB"
; OTHER INFORMATION: /note= "Disclosed in Figure 6."
; US-08-459-504B-6
Query Match 4.6%; Score 52; DB 3; Length 3624;
Best Local Similarity 47.5%; Pred. No. 0.0027;
Matches 151; Conservative 1; Mismatches 166; Indels 0; Gaps 0;
QY 100 GCGAGACTGGCGGAGAGCTKGGCTACCTGGAGATCAGCGCTGGCGTGCACCGCTTG 159
Db 181 GCGAGCTGTGGCGCGCGCGCGACCAAGTGGAGATCTTCTCTGGAGCAGCTGGAGCAG 240
QY 160 CTGGCGGACAGGAGGCTTCTGCCACCGCACTCAGCGCGCGCGCGAGCCGAGCCG 219
Db 241 CTGATCAACAGCAGATCACCGAGAACGCCCGCAACACCGCCCTGGCGCGCTGCAGGC 300
QY 220 GCGCGCGCTCGCCAGCTATCCCGAGCTGCGCTTCTGGGGGCGCTGCTGGCGCGCG 279
Db 301 CTGGGCGACAGCTTCGCGCTTACCAGCAGAGCTGGAGGACTGGCTGGAGAACCGCGAC 360
QY 280 CACTGCCTCAAGCGTGCAGAGGCGCTGCCAGCTTCCGCGAGTCCCGCAGTCCCGAGCCG 339
Db 361 GACGCCCGCAGCGCGAGCGTGTGTACACCCAGTACATCGCCCTGGAGTGGACTTCTG 420
QY 340 GAGTGTCTGGCGGACTTCCAGCGCGCGAGCGCTTCAAGTTCTCTGAGTTCGCTTACTTC 399
Db 421 AAGCCATGCCCTGTTCGCCATCCGCAACAGAGGAGTGGCGCTGCTGATGGTGTACGCC 480
QY 400 AAGGCAATAATCTCCCC 417
Db 481 CAGCGCGCAACCTGCAC 498

RESULT 8
US-08-459-444-6
; Sequence 6, Application US/08459444A
; Patent No. 6121014
; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
; Desai, Nalini M.
; Lewis, Kelly S.
; Kramer, Vance C.
; Warren, Gregory W.
; Evola, Stephen V.
; Crossland, Lyle D.
; Wright, Martha S.
; Merlin, Ellis J.
; Launis, Karen L.
; TITLE OF INVENTION: METHOD FOR PRODUCING A PLANT-OPTIMIZED
; NUCLEIC ACID CODING SEQUENCE
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 6121014artis Agribusiness Biotechnology Research, Inc.
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,444A
FILING DATE: 02-Jun-1995
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/951,715
FILING DATE: 25-SEP-1992
APPLICATION NUMBER: US 07/772,027
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Weigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: S-18805/P1/CGC1577/CIP/DIV6
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8587
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 3624 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Synthetic DNA"
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..3621
OTHER INFORMATION: /product= "Full-length, maize
optimized cryiB"
/note= "Disclosed in Figure 6."
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-08-459-444-6

Query Match 4.6%; Score 52; DB 3; Length 3624;
Best Local Similarity 47.5%; Pred. No. 0.0027;
Matches 151; Conservative 1; Mismatches 166; Indels 0; Gaps 0;
QY 100 GCGAGCAGCTGGCGCGAGAGCSTKGCTACTGTGGAGATCAGCTCGCGCTGCACCGCTTG 159
DB 181 GCGAGCTGTGGCG 240
QY 160 CTGGCGGAGAGGAGCGCTTCTGCCACCGCACTGCGAGCGCGCGCGCGCGCGCGCGCGCG 219
DB 241 CTGATCAACAGCAGATCAACCGAGAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
QY 220 GCGCGCGCGCTCGCGAGCTATCCGAGCTGCGGCTCTTCGGGGGCGCTGTGCGCGCGCG 279
DB 301 CTGGGCGAGAGCTTCGCGCGCTACCGAGAGCGCTGGAGACTGGCTGGAGAACCGCGAC 360
QY 280 CACTGCTCAAGCGCTGCAAGAGAGGCGCTTCCGAGCGCTTCCGCGAGTCCGAGCCAGCGCG 339
DB 361 GAGCGCGGACCGCGAGCGTGTGTACACCCAGTACATCGCGCTGGAGCTGAGCTTCCTG 420
QY 340 GAGTGTCTGGGAGCTTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 399
DB 421 AAGCGATGCGCTGTTCGCGATCCGCAACAGGAGTCCCGCTGCTGATGTGTAGCGCC 480
QY 400 AAGCGAAATAATCTCCCC 417
DB 481 CAGCGCGCAACCTGCAC 498

RESULT 9
US-09-053-549-7
; Sequence 7, Application US/09053549
; Patent No. 6121521
; GENERAL INFORMATION:
; APPLICANT: Desai, Nalini
; TITLE OF INVENTION: No. 6121521el Insecticidal Protein and Gene
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 6121521artis Corporation
STREET: 3054 Cornwallis Rd.
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/053,549
FILING DATE: 01-APR-1998
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Pace, Gary M.
REGISTRATION NUMBER: 40,403
REFERENCE/DOCKET NUMBER: CGC 1995
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8582
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3624 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Synthetic DNA"
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..3621
OTHER INFORMATION: /product= "Full-length, maize
optimized cryiB"
OTHER INFORMATION: /note= "Disclosed in Figure 6."
US-09-053-549-7

Query Match 4.6%; Score 52; DB 3; Length 3624;
Best Local Similarity 47.5%; Pred. No. 0.0027;
Matches 151; Conservative 1; Mismatches 166; Indels 0; Gaps 0;
QY 100 GCGAGCAGCTGGCGCGAGAGCSTKGCTACTGTGGAGATCAGCTCGCGCTGCACCGCTTG 159
DB 181 GCGAGCTGTGGCG 240
QY 160 CTGGCGGAGAGGAGCGCTTCTGCCACCGCACTGCGAGCGCGCGCGCGCGCGCGCGCGCG 219
DB 241 CTGATCAACAGCAGATCAACCGAGAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
QY 220 GCGCGCGCGCTCGCGAGCTATCCGAGCTGCGGCTCTTCGGGGGCGCTGTGCGCGCGCG 279
DB 301 CTGGGCGAGAGCTTCGCGCGCTACCGAGAGCGCTGGAGACTGGCTGGAGAACCGCGAC 360
QY 280 CACTGCTCAAGCGCTGCAAGAGAGGCGCTTCCGAGCGCTTCCGCGAGTCCGAGCCAGCGCG 339
DB 361 GAGCGCGGACCGCGAGCGTGTGTACACCCAGTACATCGCGCTGGAGCTGAGCTTCCTG 420
QY 340 GAGTGTCTGGGAGCTTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 399
DB 421 AAGCGATGCGCTGTTCGCGATCCGCAACAGGAGTCCCGCTGCTGATGTGTAGCGCC 480
QY 400 AAGCGAAATAATCTCCCC 417
DB 481 CAGCGCGCAACCTGCAC 498

RESULT 10
US-09-547-422-6
; Sequence 6, Application US/09547422
; Patent No. 6320100
; GENERAL INFORMATION:

APPLICANT: Koziel, Michael G.
Desai, Nalini M.
Lewis, Kelly S.
Kramer, Vance C.
Warren, Gregory W.
Evola, Stephen V.
Crossland, Lyle D.
Wright, Martha S.
Merlin, Ellis J.
Lauais, Karen L.
TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
INSECTICIDAL ACTIVITY IN MAIZE
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6320100artis Agribusiness Biotechnology Research, Inc.
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/547,422
FILING DATE: 11-Apr-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/459,595

FILING DATE: 02-JUN-1995

APPLICATION NUMBER: US 07/951,715

FILING DATE: 25-SEP-1992

APPLICATION NUMBER: US 07/772,027

FILING DATE: 04-OCT-1991

ATTORNEY/AGENT INFORMATION:

NAME: Meigs, J. Timothy

REGISTRATION NUMBER: 38,241

REFERENCE/DOCKET NUMBER: S-18805H

TELECOMMUNICATION INFORMATION:

TELEPHONE: (919) 541-8587

TELEFAX: (919) 541-8689

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 3624 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid

DESCRIPTION: /desc = "Synthetic DNA"

HYPOTHETICAL: NO

FEATURE:

NAME/KEY: CDS

LOCATION: 1..3621

OTHER INFORMATION: /product = "Full-length, maize

optimized cryiB"

/note = "Disclosed in Figure 6."

SEQUENCE DESCRIPTION: SEQ ID NO: 6:

US-09-547-422-6

Query Match 4.6%; Score 52; DB 4; Length 3624;

Best Local Similarity 47.5%; Pred. No. 0.0027;

Matches 151; Conservative 1; Mismatches 166; Indels 0; Gaps 0;

QY 100 GCGAGCACTGGCGGAGAGCTGGTACCTGAGATCAGCTCGCGCTGCACCGCTTG 159

DB 181 GCGAGCTGTGCGCGCGCGGCGGACCAAGTGGAGATCTTCTTGAGACAGTGGACAG 240

QY 160 CTGCGCGACAGCGAGCGCTTCTGCCACCGCAACTGCGAGCGCGCGCGAGCGCGCC 219

DB 241 CTGATCAACAGCAGATCACCAGAGACGCCCGGACACCGCGCTGGCCCGCTGCAGGGC 300

QY 220 GCGCGCGCGCTGCGCAGCTATCCGAGCTGCGCTCTTCGGGGGCTCTCGCGCGCGC 279

DB 301 CTGGCGGACAGCTTCGCGCGCTACCGAGAGCTGGAGGACTGCTGGAGAACCGCGAC 360

QY 280 CACTCCCTCAAGCGCTGCAAGCAGGGCTGCCAGCTTCGCCGCTCCAGCCAGCGGC 339

DB 361 GACGCCCGCAGCGCGAGCGTGTGTACACCGAGTACATCGGCCCTGGAGCTTGCTTG 420

QY 340 GAGGTGCTGGCGGACTTCCAGCGCGCGAGCGCTTCAAGTTCCTGCGAGTTCTGCTTACTTC 399

DB 421 AACGCCATGCCCTGTTCGCCATCGCAACAGGAGGTGCCCTGCTGATGTTGACGCC 480

QY 400 AAGCAATAATCTCCCC 417

DB 481 CAGGCCGCAACCTGCAC 498

RESULT 11

US-09-252-991A-3976

; Sequence 3976; Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 3976

; LENGTH: 942

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-3976

Query Match 4.5%; Score 50.6; DB 4; Length 942;

Best Local Similarity 46.4%; Pred. No. 0.0036;

Matches 161; Conservative 1; Mismatches 185; Indels 0; Gaps 0;

QY 26 GCTTCGCGAGCTTCCACGGGACGAGCTGATGCCCTCGAGTCGCGCTTACCGGCACGCGC 85

DB 254 GCATCGGATCTTCAACCCCGGTCAACGAACTGCCGTCGCGGCCACCGCTGCTGGSCA 313

QY 86 TGGACAAGTACAGCGCGAGACTGGCGCGAGAGCTGAGCTGAGATCAGCTGCG 145

DB 314 CGGCATCGCCCTGGCGCGCGAGACCGCAAGGACAGGCTGTTCCTGGAGACCGAGATGG 373

QY 146 GCGTCACCGCTTGTGCGCGACGAGGCTTCTGCCACCGCAACTGCGAGCGCGCGC 205

DB 374 GCACGTGCGCTTTCGCTTGGAGCGCAGAGCGGAGGTGTGCGCTTCAGCATGCGAC 433

QY 206 CGCAGCGCGAGCGCGCGCTTCGCGAGTATCCGAGCTGCGCTTTCGCGGGGCC 265

DB 434 AGCGATCCCACTTGGGAGCACTTCAGCGCGCGCGAGCTGCTCGCCGCCCTCGGCC 493

QY 266 TGCTGGCGCGCGGCACTGCGCTCAAGCGCTGCAAGAGGGCTGCGAGCTTCGCCAGT 325

DB 494 TGAAGGGTTCGACCTTTCGCATCGAGGTCTACCGCAACGCGCGCGCGCAGCTTTCGTCG 553

QY 326 CCAGCGCGCGCGAGGTGCTGGCGGAGCTTCCAGCGCGCGAGGCC 372

DB 554 GCTTGAGAGCGTCCGCGCTGTCCGCGCTGCACCCCGACCGCC 600

RESULT 12

US-09-252-991A-4046

; Sequence 4046; Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196 136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 4046
LENGTH: 1521
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4046

Query Match 4.5%; Score 50.6; DB 4; Length 1521;
Best Local Similarity 46.4%; Pred. No. 0.0042; Mismatches 185; Indels 0; Gaps 0;
Matches 161; Conservative 1;
QY 26 GTCTCGCAGCTTCCACGGGACGAGCTGATCGCGCTCGAGTCGGCCTACCGGCAAGCGC 85
DB 189 GCATCGCATCTTACCCCGGTCAACGAACTGCGGTTCGCGGCCACCGCTGCTGGCA 248
QY 86 TGGACAACTAGCGCGGAGCACTGGGCGGAGAGCTGGCTACTTGGAGATCAGCTTCG 145
DB 249 CGCCATCGCTTCCGCGCGGAGCGACGACAGAGCAGAGCTGTCTTGGAGACCAAGATGG 308
QY 146 GGCTGACCGCTGCTGCGCGGACGAGCGAGCTTCTGCCACCGCAACTGACGCGCGCGC 205
DB 309 GCACCGTGCCTTGCCTTGGAGCGGACGAGCAAGGTGTGGCTTGCAGTCAGCAGC 368
QY 206 CGCAGCCCGAGCGCGCGCGCTCGCAGCTATCCGAGTATCCGAGTGTGCGCTTTCGGGGGCC 265
DB 369 AGCCGATCCCGACCTGGGAGCACTTCAGCGCGCGCGCGAGCTGTGCGCGCTTCGCGCC 428
QY 266 TGTGCGCGCGCGCACTGCTCAAGCGCTGCAAGCAGGCGCTGCGAGCGCTTCGCCCACT 325
DB 429 TGAAGGCTCGACCTTTCCTCATCGAGTCTACGCAACGCGCGCGCGCGCACTTTCGTCG 488
QY 326 CCGAGCCCGAGCGCGAGGTGCTGGCGACTTCCAGCGCGCGAGGCC 372
DB 489 GCCTGGAGAGCGTGGCGCGCTGTCGCGCTGCAACCGCGCGCGCGCGCGCGCGCGCGCG 535

RESULT 13
US-09-105-537-30
Sequence 30, Application US/09105537A
Patent No. 6265202
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438US1
CURRENT APPLICATION NUMBER: US/09/105,537A
CURRENT FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 30
LENGTH: 13842
TYPE: DNA
ORGANISM: Streptomyces venezuelae
US-09-105-537-30

Query Match 4.4%; Score 49.8; DB 3; Length 13842;
Best Local Similarity 48.4%; Pred. No. 0.015; Mismatches 147; Indels 0; Gaps 0;
Matches 138; Conservative 0;
QY 108 CTGGCGGAGAGCGTGGCTTACCTGGAGATCAGCTTGGCGCTGCGCGCTTGTGCGCA 167

DB 9519 CCGGAGCGGCACACCGCGTCTCTTACCGGCCAGGCGAGTCAGCGCGCGCGCGCGCG 9578
QY 168 CAGGAGCGCTTTCGACCGCACTGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 227
DB 9579 CCGTGAAGCTGACAGCGGCAACCGCTTTCGCGCGCGCGCGCGCGCGCGCGCGCG 9638
QY 228 CTTGCGCAGCTATCCGAGCTGCGCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 287
DB 9639 CTTGCGCAGCTATCCGAACTGCGCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 9698
QY 288 CAGCGCTGCAAGCAGGCGCTGCGAGCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 347
DB 9699 GAGGCGCGCGCTGCTGACGAGCGCGGTACACGAGTGGCGCTTTCGCGCGCGCGCG 9758
QY 348 GCGCGACTTCCAGCG 392
DB 9759 CGCGCTTTCGCGCTGCTGAGAGCTGGGCGCATGCGCGCGCGCGCGCGCGCGCG 9803

RESULT 14
US-09-105-537-5
Sequence 5, Application US/09105537A
Patent No. 6265202
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438US1
CURRENT APPLICATION NUMBER: US/09/105,537A
CURRENT FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 5
LENGTH: 36778
TYPE: DNA
ORGANISM: Streptomyces venezuelae
US-09-105-537-5

Query Match 4.4%; Score 49.8; DB 3; Length 36778;
Best Local Similarity 48.4%; Pred. No. 0.021; Mismatches 147; Indels 0; Gaps 0;
Matches 138; Conservative 0;
QY 108 CTGGCGGAGAGCGTGGCTTACCTGGAGATCAGCTTGGCGCTGCGCGCTTGTGCGCG 167
DB 11260 CCGGAGCGGCACACCGCGTCTCTTACCGCGCGCGCGCGCGCGCGCGCGCGCG 11319
QY 168 CAGGAGCGCTTTCGCGCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 227
DB 11320 CCGTGAAGCTGACAGCGGCAACCGCTTTCGCGCGCGCGCGCGCGCGCGCGCGCG 11379
QY 228 CTTGCGCAGCTATCCGAGCTGCGCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 287
DB 11380 CTTGCGCAGCTATCCGAACTGCGCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 11439
QY 288 CAGCGCTGCAAGCAGGCGCTGCGAGCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 347
DB 11440 GAGGCGCGCGCTGCTGACGAGCGCGGTACACGAGTGGCGCTTTCGCGCGCGCGCG 11499
QY 348 GCGGACTTTCAGGCG 392
DB 11500 CGCGCTTTCGCGCTGCTGAGAGCTGGGCGCATGCGCGCGCGCGCGCGCGCGCG 11544

RESULT 15
US-09-320-878-19
Sequence 19, Application US/09320878A
Patent No. 6117659
GENERAL INFORMATION:
APPLICANT: Ashley, Gary
APPLICANT: Betlach, Melanie C.
APPLICANT: Betlach, Mary C.

```
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002120
; CURRENT APPLICATION NUMBER: US/09/320,878A
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: CIP OF 09/141,908
; EARLIER FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: CIP OF 09/073,538
; EARLIER FILING DATE: 1998-05-06
; EARLIER APPLICATION NUMBER: CIP OF 08/846,247
; EARLIER FILING DATE: 1997-04-30
; EARLIER APPLICATION NUMBER: 60/119,139
; EARLIER FILING DATE: 1999-02-08
; EARLIER APPLICATION NUMBER: 60/100,880
; EARLIER FILING DATE: 1998-09-22
; EARLIER APPLICATION NUMBER: 60/087,080
; EARLIER FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 38506
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-320-878-19

Query Match      4.4%; Score 49.8; DB 3; Length 38506;
Best Local Similarity 48.4%; Pred. No. 0.022;
Matches 138; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

QY 108 CTGGGCGGAGGCTKGGCTACCTGGAGATCAGCTCGGCTGCACCGCTTGTCTGGCGGA 167
Db 9402 CCGGGACGGACACACCGCGTCTTTCACCGGCGAGGGCAGTCAGCGCCCGCGCGCG 9461

QY 168 CAGCGAGGCTTCTGCCACCGCAACTCGACGGCGCGCGCGCGAGCCCGCGCGCG 227
Db 9462 CCGTGAGCTGTACGACCGGACACCGCGTCTTTCGCGCGCGCTCGACGAGATCTGCGCCCA 9521

QY 228 CCTCGCCAGCTATCCGAGCTGGCTCTTTCGGGGGCTCTCGCGCGCGCGCAGTGCCT 287
Db 9522 CCTCGAGGCTACCTCGAACTGCCCTGCTCGACGTGATGTTTCGCGCGCGAGGGCAGCG 9581

QY 288 CAAGCGCTGCAAGCAGGGGCTGCGAGCTTCCGCCAGTCCCGAGCCAGCGCGAGGTGCT 347
Db 9582 GGAGGCGCGCTGCTCGACGAGCGCGGTACGCGAGTGGCGCTGTTCCGCTTGAGGT 9641

QY 348 GGCGGACTTCCAGCGCGCGAGCCCTACAGTTCCTGCACTTCCG 392
Db 9642 CGCGCTTTCGGGCTCGTCGAGAGCTGGGGCATCGCGCGCGCGC 9686
```

Search completed: January 30, 2004, 18:32:30
Job time : 69.9336 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2004, 07:55:54 ; Search time 2380.16 Seconds
(without alignments)
11579.600 Million cell updates/sec

Title: US-09-729-674-1_COPY_132_1265

Perfect score: 1334

Sequence: 1 999gcgcacatacgaacg.....tgaactggaggagaccagc 1134

Scoring table:

IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:

1: em_estba:

2: em_esthum:

3: em_estin:

4: em_estm:

5: em_estov:

6: em_estpl:

7: em_estro:

8: em_hc:

9: gb_est1:

10: gb_est2:

11: gb_hc:

12: gb_est3:

13: gb_est4:

14: gb_est5:

15: em_estfun:

16: em_estom:

17: em_gss_hum:

18: em_gss_inv:

19: em_gss_pln:

20: em_gss_vrt:

21: em_gss_fun:

22: em_gss_mam:

23: em_gss_mus:

24: em_gss_pro:

25: em_gss_rod:

26: em_gss_phg:

27: em_gss_vrl:

28: gb_gss1:

29: gb_gss2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	981.2	85.5	1201	9	AL545483
2	920.4	81.2	1201	9	AL556976
3	908.2	80.1	1201	9	AL514488
4	897.2	79.1	1656	11	AK047506 Mus muscu

5	893.8	78.8	1201	13	BX425905
6	889	78.4	1148	13	BU902192
7	886	78.1	1652	11	AK017797
8	877.8	77.4	1075	9	AL546910
9	873.8	77.1	1201	9	AL517455
10	865.6	76.3	1201	9	AL541167
11	811.6	71.6	1201	9	AL554750
12	807.8	71.2	1201	9	AL543090
13	801	70.6	846	13	BU838140
14	784.4	69.2	1109	12	BM550012
15	771.4	68.0	774	9	AL598447
16	769	67.8	939	13	BX386576
17	768.8	67.8	1201	9	AL558186
18	758.2	66.9	1152	9	AL550966
19	747.4	65.9	1200	13	BU902214
20	741.8	65.4	1040	9	AL558966
21	727.6	64.2	944	13	BU856662
22	720.6	63.5	1430	11	AK007624
23	718	63.3	1201	13	BX437872
24	711.8	62.8	945	13	BQ677531
25	701.6	61.9	1201	13	BX387779
26	698	61.6	846	13	BU189067
27	697.8	61.5	1083	12	BM920687
28	697.6	61.5	713	10	BE746070
29	692.2	61.0	846	12	B1253401
30	691.6	61.0	887	13	BQ718961
31	691.4	61.0	899	13	BQ437797
32	683.2	60.2	904	13	BU931550
33	678.4	59.8	878	13	BU183287
34	677.2	59.7	962	14	CD516948
35	674.8	59.5	949	10	BG323659
36	674.4	59.5	1103	13	BQ233085
37	669.8	59.1	934	12	BM453316
38	668	58.9	827	12	B1760650
39	665.2	58.7	814	13	BQ953454
40	658.8	58.1	844	10	BE744854
41	657	57.9	875	14	CD359485
42	655.2	57.8	912	13	BU175306
43	641.2	56.5	1201	13	BX460044
44	638.4	56.3	770	10	BE783622
45	635	56.0	782	10	BE877856

ALIGNMENTS

RESULT 1
AL545483 1201 bp mRNA linear EST 31-MAY-2003
LOCUS
DEFINITION AL545483 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CSODI015YK22 5-PRIME, mRNA sequence.
ACCESSION AL545483
VERSION AL545483.2 GI:31267318
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT On Feb 15, 2001 this sequence version replaced gi:12877964.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 2233.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSODI015BP11Q1P1&cluster=2233.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :

```
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue Genoscope sequence ID : CS0D1015BF11Q1.
FEATURES
Source
1. 1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1015YK22"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      256 a   363 c   321 g   242 t   19 others
ORIGIN
Query Match      86.5%; Score 981.2; DB 9; Length 1201;
Best Local Similarity 97.2%; Pred. No. 1.5e-227;
Matches 999; Conservative 17; Mismatches 11; Indels 1; Gaps 1;
QY 8 CCCAATAGCAAGCTTACAGCTTCCCGAGCTTCCACGGGAGGAGCTGATGCCGCTCGAGT 67
DB 180 CCCAATAGCAAGCTTACAGCTTCCCGAGCTTCCACGGGAGGAGCTGATGCCGCTCGAGT 239
QY 58 CGGCTACGGACCGCTGACAGTACAGGGCGGACACTGGCGGAGAGCTKGCT 127
DB 240 CGGCTACGGACCGCTGACAGTACAGGGCGGACACTGGCGGAGAGCTGGGCT 299
QY 128 ACCTGGAGATCAGCTCGGGCTGACCCGCTTGTCTGGCGGACGAGCGCTTCTGCCACC 187
DB 300 ACCTGGAGATCAGCTCGGGCTGACCCGCTTGTCTGGCGGACGAGC-AGGCTTCTGCCACC 358
QY 188 GCAATGACAGCCCGCGGACGCGGAGCGCGCGCGGCTCGCCAGCTATCCGAGC 247
DB 359 GCAATGACAGCCCGCGGACGCGGAGCGCGCGCGGCTCGCCAGCTATCCGAGC 418
QY 248 TCGGCTCTTGGGGGCTGTCTGGCGCGGCGACTGCTCAAGCGTGCAGGAGGCGC 307
DB 419 TCGGCTCTTGGGGGCTGTCTGGCGCGGCGACTGCTCAAGCGTGCAGGAGGCGC 478
QY 308 TGCCAGCTTCCGAGTCCGAGCCGAGCGCGGAGTGTCTGGCGACTTCCAGCGCGCG 367
DB 479 TGCCAGCTTCCGAGTCCGAGCCGAGCGCGGAGTGTCTGGCGACTTCCAGCGCGCG 538
QY 368 AGCCCTACAGTTCCTGAGTTCGTTTCAAGGCAAAATCTCCCAAGCCATCG 427
DB 539 AGCCCTACAGTTCCTGAGTTCGTTTCAAGGCAAAATCTCCCAAGCCATCG 598
QY 428 CGGCTGCTCACACCTTTCTACTGAGCATCTCTGATGACGAAATGATGAAGGAGCACTGG 487
DB 599 CGGCTGCTCACACCTTTCTACTGAGCATCTCTGATGATGAATGATGAAGGAGCACTGG 658
QY 488 CATATTAAAGAGCTGCTGTGTCGGAGGACTACATTAAGAGCTTGAAGACCAAGTCAT 547
DB 659 CATATTAAAGAGCTGCTGTGTCGGAGGACTACATTAAGAGCTTGAAGACCAAGTCAT 718
QY 548 ATGAAGCTGTTCATCCGAGCAGTCCGCGGATACACCGGTGAGAACTGGAGAACATCCA 607
DB 719 ATGAAGCTGTTCATCCGAGCAGTCCGCGGATACACCGGTGAGAACTGGAGAACATCCA 778
QY 608 TCACAGACATGAGTGGCCCTTCCCGACTTCTTCAAGCCCTTTTACGAGTGTCTCGAG 667
DB 779 TCACAGACATGAGTGGCCCTTCCCGACTTCTTCAAGCCCTTTTACGAGTGTCTCGAG 838
QY 668 CTTGCGAGGTTCCAGGGAGATCAGGACTTTCAGGATTTCTACTTTCCATACAGATC 727
DB 839 CTTGCGAGGTTCCAGGGAGATCAGGACTTTCAGGATTTCTACTTTCCATACAGATC 898
QY 728 ATTATGTAGAAGTCTTGAATCAAAATACAGTGTGAAGAGAACTCCACCCCAAGTTATG 787
DB 899 ATTATGTAGAAGTCTTGAATCAAAATACAGTGTGAAGAGAACTCCACCCCAAGTTATG 958
```

```
QY 788 GAGGCTATCGGTTGAGAAATTTGGCTACCATGATCATTTGAGTTGGCTATT 847
DB 959 GAGGCTATCGGTTGAGAAATTTGGCTACCATGATCATTTGAGTTGGCTATT 1018
QY 848 ATAAAGTTGAACGACCTGAAGAATGAGCCCTCTGTGAGTACGCTATCTGCTTTGATC 907
DB 1019 ATAAAGTTGAACGACCTGAAGAATGAGCCCTCTGTGAGTACGCTATCTGCTTTGATC 1078
QY 908 AGAATGACAAGTTCATGTCAGCAGAACCTGCTGTTATACCAGTACCACAGGACACTTGGG 967
DB 1079 AGAATGACAAGTTCATGTCAGCAGAACCTGCTGTTATACCAGTACCACAGGACACTTGGG 1138
QY 968 GCCTCTCGATGAGCACTTCCAGCCGACACTGAGAGCACTTCAAGTCTTTTAATGATC 1025
DB 1139 CCTCTCGATGAGCACTTCCAGCCGACACTGAGTCTTCAAGTCTTTTAATGATC 1196

RESULT 2
AL556976 1201 bp mRNA linear EST 31-MAY-2003
LOCUS AL556976 Homo sapiens T CELLS (JURKAT CELL LINE) Homo sapiens cDNA
DEFINITION clone CS0DH005YI01 5-PRIME, mRNA sequence.
ACCESSION AL556976
VERSION AL556976.2 GI:31278776
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1. (bases 1 to 1201)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT On Feb 15, 2001 this sequence version replaced gi:12900137.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
2233.f For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DH005YI01&cluster=2233.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue Genoscope sequence ID : CS0DH005YI01Q1.
FEATURES
Source
1. 1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DH005YI01"
/tissue_type="T CELLS (JURKAT CELL LINE)"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE)"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
BASE COUNT      272 a   339 c   320 g   234 t   36 others
ORIGIN
Query Match      81.2%; Score 920.4; DB 9; Length 1201;
Best Local Similarity 97.8%; Pred. No. 8.9e-213;
Matches 943; Conservative 9; Mismatches 10; Indels 2; Gaps 2;
QY 9 CCAATAGCAAGCTTACAGCTTCCGAGCTTCCACGGGAGGAGCTGATGCCCTCGAGTC 68
DB 131 CCAATAGCAAGCTTACAGCTTCCGAGCTTCCACGGGAGGAGCTGATGCCCTCGAGTC 190
QY 69 GGCTTACCGGCGAGCTGAGCAAGTACAGCGCGGAGCACTGGCGGAGAGGCTGGGCTA 128
```



```
Db 191 GGCCTACCGGACGCGCTGGACAGTACAGCGCGAGCAGCTGKCGGAGAGCGTGGGCTA 250
Qy 129 CTGGAGATAGCTGGCTGGACCGCTTCTGCGGACAGCGAGCGCTTCTGCCACCG 188
Db 251 CTGGAGATAGCTGGCTGGACCGCTTCTGCGGACAGC-AGGCTTCTGCCACCG 309
Qy 189 CAAGTGCAGCGCGCGCGCGAGCGCGCGCGCGCGCGCTGCGGAGCTATCCGAGCT 248
Db 310 CAAGTGCAGCGCGCGCGCGAGCGCGCGCGCGCGCGCTGCGGAGCTATCCGAGCT 369
Qy 249 GCGCTCTTTCGGGCGCTGCTGCGCGCGCGCGCTCAAGCGCTGCAAGCAGGCGCT 308
Db 370 GCGCTCTTTCGGGCGCTGCTGCGCGCGCGCGCTCAAGCGCTGCAAGCAGGCGCT 429
Qy 309 GCGAGCTTTCGGGCGCTGCTGCGCGCGCGCGCTGCGGAGCTATCCGAGCGCGA 368
Db 430 GCGAGCTTTCGGGCGCTGCTGCGCGCGCGCGCTGCGGAGCTATCCGAGCGCGA 489
Qy 369 GCGCTACAGCTTCTGAGCTTCTGCTTACTTCAAGCGCAATTAATCTCCCGAAGCGATCGC 428
Db 490 GCGCTACAGCTTCTGAGCTTCTGCTTACTTCAAGCGCAATTAATCTCCCGAAGCGATCGC 549
Qy 429 GCGCTACAGCTTCTGAGCTTCTGCTTACTTCAAGCGCAATTAATCTCCCGAAGCGATCGC 488
Db 550 GCGCTACAGCTTCTGAGCTTCTGCTTACTTCAAGCGCAATTAATCTCCCGAAGCGATCGC 609
Qy 489 ATATTATAAGAGCTGCTGCTGCGGAGGACTACATTAAAGACTTGAAGAGCGATCGA 548
Db 610 ATATTATAAGAGCTGCTGCTGCGGAGGACTACATTAAAGACTTGAAGAGCGATCGA 669
Qy 549 TGAAGCGCTTTCATCGAGAGCTGCGGCGATCAACGCTGAGAGCTGAGAGCAATCCAT 608
Db 670 TGAAGCGCTTTCATCGAGAGCTGCGGCGATCAACGCTGAGAGCTGAGAGCAATCCAT 729
Qy 609 CACAGACATGAGCTGCGGCTTCCGAGCTTCTCAAGCGCTTTCAGAGCTTTCAGAGCTTCCGAGC 668
Db 730 CACAGACATGAGCTGCGGCTTCCGAGCTTCTCAAGCGCTTTCAGAGCTTTCAGAGCTTCCGAGC 789
Qy 669 CTGCGAGGTTCCAGGAGATCAAGGACTTCAAGGATTTCTACTTCCATAGCAGATCA 728
Db 790 CTGCGAGGTTCCAGGAGATCAAGGACTTCAAGGATTTCTACTTCCATAGCAGATCA 849
Qy 729 TTAGTGTAGAGTTCTGGAATGCAAAATACAGTGTGAAGAGAACTCAACCCAGTTATAGG 788
Db 850 TTAGTGTAGAGTTCTGGAATGCAAAATACAGTGTGAAGAGAACTCAACCCAGTTATAGG 909
Qy 789 AGGCTATCCGTTGAGAAATTTCTGCTACCATGTATCATTTACTTGCAGTTTGCCTATTA 848
Db 910 AGGCTATCCGTTGAGAAATTTCTGCTACCATGTATCATTTACTTGCAGTTTGCCTATTA 969
Qy 849 TAAATTGAAGCCTGAAGATGACGCGCGCTGTCAGTCACTATCTCTTTGATCA 908
Db 970 TAAATTGAAGCCTGAAGATGACGCGCGCTGTCAGTCACTATCTCTTTGATCA 1029
Qy 909 GAATGCAAGGTCATGACAGAACTGCTGTGATTAACAGTACAGGAGCACTTGGG 968
Db 1030 GAATGCAAGGTCATGACAGAACTGCTGTGATTA-CACTACCAVAGAGCACTTGGG 1089
Qy 969 CTTG 972
Db 1089 YYVC 1092

RESULT 3
AL514488
LOCUS
DEFINITION
AL514488 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
ACCESSION
AL514488
VERSION
AL514488.2 GI:30464373
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished

On Feb 13, 2001 this sequence version replaced gi:12777992.

Contact: Genoscope

Genoscope - Centre National de Sequencage

Bp 191 91006 EVRY cedex - France

Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 2233.f For

more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CLOB003ZF09RPI&cluster=2233.f. Contact :

Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ invitrogen Corporation 1600

Paradise Avenue Genoscope sequence ID : CLOB003ZF09RPI.

Location/Qualifiers

1..1201

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CLOB003ZF09"

/tissue_type="NEUROBLASTOMA"

/clone_lib="Homo sapiens NEUROBLASTOMA"

/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed

with a NotI-oligo(dT) primer. Five prime end enriched

double-strand cDNA was digested with Not I and cloned into

the Not I and EcoRV sites of the pCMVSPORT 6 vector.

Library was not normalized."

BASE COUNT 274 a 299 c 324 g 246 t 58 others

ORIGIN

Query Match 80.1%; Score 908.2; DB 9; Length 1201;

Best Local Similarity 93.7%; Pred. No. 8.2e-210;

Matches 953; Conservative 31; Mismatches 29; Indels 4; Gaps 4;

Qy 2 GCGCGGCCAATACGAACGCTACAGCTTCCGAGCTTCCACGGGACGAGTGTGCGC 61

Db 132 GCGCGGCCAATACGAACGCTACAGCTTCCGAGCTTCCACGGGACGAGTGTGCGC 191

Qy 62 TCGAGTGGCTTACCGGACGCGCTGACAGTACAGGACGAGGCGGACGAGCGGAG 121

Db 192 TCGAGTGGCTTACCGGACGCGCTGACAGTACAGGACGAGGCGGACGAGCGG 251

Qy 122 TGGCTACCTGGAGATCAGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTTCT 181

Db 252 TGGCTACCTGGAGATCAGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTTCT 311

Qy 182 GCGACCGCAACTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTATC 241

Db 312 GCGACCGCAACTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTATC 371

Qy 242 CCGAGCTGCGGCTTCTCGGGGCGCTGCTGCGGCGCGCGCGCGCGCGCGCGCGCG 301

Db 372 CCGAGCTGCGGCTTCTCGGGGCGCTGCTGCGGCGCGCGCGCGCGCGCGCGCGCG 429

Qy 302 AGGCGCTGCGGCTTCTCGGGGCGCTGCTGCGGCGCGCGCGCGCGCGCGCGCGCG 361

Db 430 AGGCGCTGCGGCTTCTCGGGGCGCTGCTGCGGCGCGCGCGCGCGCGCGCGCGCG 489

Qy 362 GCGCGGAGCGCTTCTCGAGTTCCTGAGTTCCTGAGTTCCTGAGTTCCTGAGTTC 421

Db 490 GCGCGGAGCGCTTCTCGAGTTCCTGAGTTCCTGAGTTCCTGAGTTCCTGAGTTC 549

Qy 422 CCAATCGCGCTGCTCAACCTTTTACTGAAGATCTCTGATGACGAAATGATGAAG 481

Db 550 CCAATCGCGCTGCTCAACCTTTTACTGAAGATCTCTGATGACGAAATGATGAAG 609

Qy 482 ACATGGCATATTATAGAGCGCTGCGGCGGAGCTACATTAAAGACCTGGAAACCA 541


```
/note="unnamed protein product; cartilage associated
protein (MGI:1891221, GB|NM_019922, evidence: BLASTN,
99%, match=1655)
putative"
/codon_start=1
/protein_id="BAC33076.1"
/db_xref="GI:26338810"
/translation="MGRPSPTAALLVLLVCVCAPTPGRGVYRSFRSPRDELMPLE
SAVRLHQVSGHWAESVGLVLSRLHLLRDSFAFCHRNCSAATPAPAPAGPASH
AEURLFQSVLRQAQCLKRCQGLPAFRQSPRSVLADFQREPYFLQFAYFKANDL
PKAIAAHTYLLKHPDDMMKRMNYYKSLPQAEHDIKDLTKSYESLFFVRAVYNG
ENWRTSLSDMELALPDLKAFYECIAACEGSRDKDFKDFYLSIAHYHVECKLRIC
EELTLPIVIGVPEKFWATWYHLYQFAYKLNKNAAPCAVSLLPDSDRWMOONL
VYQYHREDKXWLSDEHFQPRPEAVQFENWTLQKLYDFPAQHLLMDDEGEVVEYVD
LLETEESA"
BASE COUNT 341 a 498 c 448 g 369 t
ORIGIN
Query Match 79.1%; Score 897.2; DB 11; Length 1656;
Best Local Similarity 87.0%; Pred. No. 4.6e-207;
Matches 983; Conservative 2; Mismatches 145; Indels 0; Gaps 0;
QY 1 GGGGCGCCCAATACGAACGCTACAGTTCCGAGCTTCCCGGACGAGCTGATGCCG 60
DB 136 GGGCGCGGGAGTACGAGCGCTACAGCTTCCGAGCTTCCCGGGACGAGCTGATGCCG 195
QY 61 CTGAGTCGCGCTACCGGACGCGCTGACAGTACAGCGGCGAGCACTGGCGCGAGAC 120
DB 196 CTGAGTCGCGCTACCGGACGCGCTGACAGTACAGCGGCGAGCACTGGCGCGAGAC 255
QY 121 STKGCTACTGAGATGACGCTCGGCTGACCGCTGCTGCGCGACAGCGGCTTC 180
DB 256 GTGGCTTACCTGAGGTGAGCTTGGGCTGACCGTCTGCTGCGCGACAGCGGCTTC 315
QY 181 TGCACCGCAACTGACGCGCGCGCGCGAGCCCGCGCGCGCGCGCGCGCGCGCTG 240
DB 316 TGCACCGCAACTGACGCGCGCGCGCGAGCCCGCGCGCGCGCGCGCGCGCGCGCTG 375
QY 241 CCGGAGTCGCGCTCTTTCGGGGGCGCTGCTGCGCGCGCGCGCGCTCAAGGCTCAAG 300
DB 376 GCGGAATACGCTCTTTCGGAGGTGCTGCGCGCGCGCGCGCTCAAGGCTCAAG 435
QY 301 CAGGCTGCGCGCTTCCCGGAGTCCGCGAGTCCGCGAGCGCGGAGGTGTCGCGAGTCCAG 360
DB 436 CAGGCTGCGCGCTTCCCGGAGTCCGCGAGTCCGCGAGCGCGGCTCAAGTCCAG 495
QY 361 CGCGGAGCGCTTACAGTCTTCGAGTTCGCTTCTCAAGGCAATATCTCCGCCAA 420
DB 496 CAGGCGAGCGCTTACAGTCTTCGAGTTCGCTTCTCAAGGCAATATCTCCGCCAA 555
QY 421 GCATCGCGCTGCTCACACCTTCTTACTGAAGCATCTCTGATGACGAAATGATGAAGG 480
DB 556 GCATCGCTCGGCTCACACCTATCTCTGAAGCATCTCTGATGACGAAATGATGAAGG 515
QY 481 AACATGCGATATTATGAAGCTGCTGTCGCGAGGACTACATTAAGACCTGGAAC 540
DB 616 AACATGAGTATTATGAAGCTGCTGTCGCGAGGAGCCACATTAAGACCTGGAAC 675
QY 541 AAGTCATATGAAGCTGTTTATCCGAGGAGTCCGCGGATACAAACGCTGAGAACTGGAGA 600
DB 676 AAGTCATGAGAGGCTGTTTATCCGAGGAGTCCGCGGATACAAACGCTGAGAACTGGAGA 735
QY 601 ACATCCATCACAGATGAGCTTGGGCTTCCGAGTCTTCAAGACCTTTTACAGTGT 660
DB 736 ACCTCCATTTCCGAGATGAGCTTGGGCTTCCGAGTCTTCAAGACCTTTTACAGTGT 795
QY 661 CTGCGAGCTTCCGAGGTTTCCAGGAGATCAAGGCTTCAAGGATTTTACCTTTCCATA 720
DB 796 CTGCTGCTTCCGAGGTTTCCGAGGAGATCAAGGCTTCAAGGATTTTACCTTTCCATA 855
QY 721 GCAGATCATTTATGAAGTTCTGGAATGCAAAATACAGTGTGAAGAACTTCAACCCA 780
DB 856 GCAGATCACTATGTGAAGTTCTGGAATGCAAAATACAGTGTGAAGAACTTCAACCCA 915
```

```
QY 781 GTTATAGGAGGCTATCCGGTTGAGAAATTGTGGCTACATGTATCATTTAGTTCAGTTT 840
DB 916 GTCATAGGAGGCTATCCCGTGGAGAAATTGTGGGAGCAATGTACCATTTTACAGTTT 975
QY 841 GCGTATTATAGTTTGAAGCACTGGAAGATGAGAGCCCTGTGCAAGTACAGTATCTGTC 900
DB 976 GCGTATTATAGTTTGAAGCACTGGAAGATGAGAGCCCTGTGCAAGTACAGTATCTGTC 1035
QY 901 TTTGATCAGATGACAAAGTTCATGACAGAGCACTGTTGTTATACAGTACACAGGAC 960
DB 1036 TTTGACAGAGTACAGAGGTTATGCAACAGAACCTGTTGTTATCAGTACACAGGAC 1095
QY 961 ACTTGGGCGCTTCTCGATGAGCACTTCCAGCCAGCACTGGAAGCAGTTTCTTTTAA 1020
DB 1096 AAGTGGGCGCTTCTCGATGAGCACTTCCAGCCAGCACTGGAAGCAGTTTCTTTTAA 1155
QY 1021 GTGACCACTCCAGAGGAGCTGTATGACTTCTTGAAGAAATATATGATGATGAT 1080
DB 1156 GTGACCACTCCAGAGGAGGAACTGTACGACTTGTGAGAACCACTTAATGATGAG 1215
QY 1081 CAGGAGAGGTTGTGGAATATGTGATGACCTTCTTGAAGTGGAGGAGAC 1130
DB 1216 CAGGAGAGGTTGTGAGTATGTGAGCACTTGTGAGAGCGAGAGTC 1265
RESULT 5
BX425905 1201 bp mRNA linear EST 15-MAY-2003
BX425905 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
CS0DA004YM07 5-PRIME, mRNA sequence.
ACCESSION BX425905
VERSION BX425905.1 GI:30774508
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 2233.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0AA004AG04QPI&cluster=2233.f. Contact :
Peng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0AA004AG04QPI.
Location/Qualifiers
1. 1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DA004YM07"
/tissue_type="NEUROBLASTOMA"
/clone_lib="Homo sapiens NEUROBLASTOMA"
/note="vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with NotI and cloned into
the NotI and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
BASE COUNT 250 a 334 c 314 g 248 t 55 others
ORIGIN
```

```
Query Match 78.8%; Score 893.8; DB 13; Length 1201;
Best Local Similarity 93.4%; Pred. No. 2.6e-206;
Matches 944; Conservative 28; Mismatches 31; Indels 8; Gaps 4;
```


FEATURES		Location/Qualifiers		BASE COUNT		ORIGIN	
source		1..1652		338 a 503 c 445 g 366 t			
/organism="Mus musculus"							
/mol_type="mRNA"							
/strain="C57BL/6J"							
/db_xref="PANTOM DB:5730529N23"							
/db_xref="MGI:1898560"							
/db_xref="taxon:10090"							
/clone="5730529N23"							
/issue_type="whole body"							
/clone_lib="RIKEN full-length enriched mouse cDNA library"							
/dev_stage="8 days embryo"							
77..1279							
CDS							
/note="unnamed protein product; cartilage associated protein (MGI:1891221, GB NM_019922, evidence: BLASTN, 99%, match=1655)							
putative"							
/codon_start=1							
/protein_id="BAB30938.1"							
/db_xref="GI:12857227"							
/db_xref="MGI:1891221"							
/translation="MGPSPTAALLVLCVCAHPGGOVERYSFSPRDELMPLE SAYRHALQYGEHWAESVGVLEVSRLRLHLLRBRQAFCHENCSSAAFPAPAPGASH ALRFLFGVLRQAQCKKQSLPAFROSQSRSLADPQOREFYKLFQPAIFRANGL PRAIAAHTYLLKHKDDDEMRKNMEYYKSLPGEADHIKOLETKSYSLFVRAVRANGL ENWRTSIDMELALPDFKAFYECIAACEGVAETKPKDFYLSIADHYVELECKIRC BETLTPVIGGYPVEFVATMYHYLQFAYKLANDKNAAPCAVSYLLDFQSDRVNQNL VYQYHRDKWGLSDHFPQPEAVQFENVVTLQLKELYDFAQEHMLDDDEGEVVEYVD LLETFESA"							
Query Match		78.1%; Score 886; DB 11; Length 1652;					
Best Local Similarity		86.4%; Pred. No. 2.4e-204;					
Matches		97%; Conservative 2; Mismatches 152; Indels 0; Gaps 0;					
QY	1	GGGCGGCCCATATAGAACGCTACAGCTTCGACGCTTCCACGGAGAGCTGATGCGG	60				
DB	143	GGGCGGGCAGTAGAGCGCTACAGCTTCGACGCTTCCGCGGGAGAGCTGATGCGG	202				
QY	61	CTCGAGTCGGCTACCGGACCGCGCTGGACAGTACAGCGCGGAGCACTGGCGGAGAGC	120				
DB	203	CTCGAGTCGGCTACCGGACCGCGCTGCACAGTACAGCGCGGAGCACTGGCGGAGAGC	262				
QY	121	STKGCTACCTGGAGATCAGCTCGCGCTGCAACCGCTTGTGCGCGAGAGAGCGCTTC	180				
DB	263	GTGGGTACCTGAGGTGAGCTCGCGCTGCAACCGCTTGTGCGCGAGAGAGCGCTTC	322				
QY	181	TGCCACCGCACTGACGCGCGCGCGCGAGCGCGCGCGCGCGCTTCGCGAGCTAT	240				
DB	323	TGCCACCGCACTGACGCGCGCGCGCGAGCGCGCGCGCGCGCTTCGCGAGCTAT	382				
QY	241	CCGAGCTCGGCTCTTGGGGGCTGTGTGCGCGGCGGCACTGCTCAAGCGCTGCAAG	300				
DB	383	GCGCAACTACGCGCTTTCGCGAGCGCTGCTGCGCGCGGCGAGCTCAAGCGCTGCAAG	442				
QY	301	CAGGCTCGCGAGCTTCGCGAGTCCAGCGCGCGCGCGAGCTGCTGCGAGCTTCAG	360				
DB	443	CAGGCTCGCGAGCTTCGCGAGTCCAGCGCGCGCGAGCTGCTGCGAGCTTCAG	502				
QY	361	CCGCGGAGCGCTTACAGTTCCTGCGAGTTCGCTTACTTCAAGCGCAAAATCTCCCGAAA	420				
DB	503	CAGCGGAGCGCTTACAGTTCCTGCGAGTTCGCTTACTTCAAGCGCAAAATCTCCCGAAA	562				
QY	421	GCATGCGCGCTGCTCACACCTTTCTAGTGAAGCATCTGTATGAGCAATATGATGAGAGG	480				
DB	563	GCATGCGCGCTGCTCACACCTTTCTAGTGAAGCATCTGTATGAGCAATATGATGAGAGG	622				
QY	481	AACATGGCATATTATTAAGAGCGCTGCTGCTGCGAGGACTACATTAAGAGCTTGAAACC	540				
DB	623	AACATGGCATATTATTAAGAGCGCTGCTGCTGCGAGGACTACATTAAGAGCTTGAAACC	682				
QY	541	AAGTCATATGAAGCGCTGTTCAATCCGAGCAGTTCGCGGGCATACACGGGTGAGAACTGGAGA	600				
DB	583	AGTCTGACGAGAGCTGTTTCTCGTGGGTGGGGCTACACGGGAGAGCTGGAGA	742				
QY	601	ACATCCATCAGACATGAGAGCTGGCCCTTCCGACTTCTTCAAAGCCCTTTACAGAGTG	660				
DB	743	ACGTCCATTTCCGACATGAGAGCTGGCCCTTCCGACTTCTTCAAAGCCCTTTACAGAGTG	802				
QY	661	CTCGAGCTCGGAGGTTTCCAGGAGATCAAGGACTTCAAGGATTTCTACCTTTCCCAT	720				
DB	803	CTCGCTGCTCGAGGGGTCGCGGAGATCAAGGACTTCTACCTTTCCCAT	862				
QY	721	GCAGATCATTTGTAGAAGTTCTGGAATGCAAAATACAGTGTGAAGAGAACCTCACCCCA	780				
DB	863	GCAGATCATTTGTAGAAGTTCTGGAATGCAAAATACAGTGTGAAGAGAACCTCACCCCA	922				
QY	781	GTATAGGAGGCTATCCGCTTGGAGAAATTTGGGCTACCATGATATCATTTACGTTT	840				
DB	923	GTATAGGAGGCTATCCGCTTGGAGAAATTTGGGCTACCATGATATCATTTACGTTT	982				
QY	841	GCCTATTATAAGTTGAACGACCTGAAGAAATGAGGCCCCCTGTGCACTGAGCTATCTGCTC	900				
DB	983	GCCTATTATAAGTTGAATGATCTGAAGAAATGAGGCCCCCTGTGCACTGAGCTATCTGCTC	1042				
QY	901	TTTGATCAGATCAGACAGCTCATGAGCAGAGACCTGGTGTATTACAGTACACAGGAGC	960				
DB	1043	TTTGATCAGATCAGACAGCTCATGAGCAGAGACCTGGTGTATTACAGTACACAGGAGC	1102				
QY	961	ACTTGGGGCTCTCGGATGAGCACTTCCAGCCGACAGCTTGAAGCAGTTCAGTTCCTTTAT	1020				
DB	1103	AAGTGGGGCTCTCGGATGAGCACTTCCAGCCGACAGCTTGAAGCAGTTCAGTTCCTTTAT	1162				
QY	1021	GTGACACACTCCAGAGGAGCTGTATGACTTGTGAGGAAATATATGATGATGAT	1080				
DB	1163	GTGACACACTCCAGAGGAGCTGTATGACTTGTGAGGAAATATATGATGATGAT	1222				
QY	1081	GAGGAGAGAGTTTGTGAATATGTGGATGACCTTCTTGAAGCTGGAGAGAGC	1130				
DB	1223	GAGGAGAGAGTTTGTGAATATGTGGATGACCTTCTTGAAGCTGGAGAGAGC	1272				
RESULT 8		1075 bp mRNA linear EST 31-MAY-2003					
LOCUS		AL546910 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA					
DEFINITION		clone CS0D1032YH24 5-PRIME, mRNA sequence.					
ACCESSION		AL546910					
VERSION		AL546910.2					
KEYWORDS		EST.					
SOURCE		Homo sapiens (human)					
ORGANISM		Homo sapiens					
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
TITLE		Li, W.B., Gruber, C., Jessee, J. and Polayes, D.					
JOURNAL		Full-length cDNA libraries and normalization					
COMMENT		Unpublished					
		On Feb 15, 2001 this sequence version replaced gi:12880487.					
		Contact: Genoscope					
		Genoscope - Centre National de Sequenage					
		BP 191 91006 EVRY cedex - France					
		Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr					
		Library was constructed by Life Technologies, a division of					
		Invitrogen. This sequence belongs to sequence cluster 2233.f For					
		more information about this cluster, see					
		http://www.genoscope.cns.fr/					
		cgi-bin/cluster.cgi?seq=CS0D1032YH24P1&cluster=2233.f. Contact :					
		Feng Liang Email : fliang@lifetech.com URL :					
		http://fulllength.invitrogen.com/Invitrogen Corporation 1600					
		Faraday Avenue Genoscope sequence ID : CS0D1032YH24P1.					
FEATURES		Location/Qualifiers					
source		1..1075					
		/organism="Homo sapiens"					
		/mol_type="mRNA"					


```

Db 372 GCGCCTCTTCGGGGGCTGCTGGCGCGCGCACTGCCTCAAGCGTGAAGCAGGGCT 431
QY 309 GCAGCGCTTCGGCGAGTCCAGAGCCAGCGAGGCTGGCGGAGTTCAGCGCGCGGA 368
Db 432 GCAGCGCTTCGGCGAGTCCAGAGCCAGCGAGGCTGGCGGAGTTCAGCGCGCGGA 491
QY 369 GCCTCAAGTTCCTGCAGTTCCTTACTTCAAGGCAAAATATCTCCCAAGCCATGC 428
Db 492 GCCTCAAGTTCCTGCAGTTCCTTACTTCAAGGCAAAATATCTCCCAAGCCATGC 551
QY 429 CGCTGCTCACACCTTCTACTGAGCATCTGATGAGCAATGATCAAGAGCAATGCG 488
Db 552 CGCTGCTCACACCTTCTACTGAGCATCTGATGAGCAATGATCAAGAGCAATGCG 611
QY 489 ATATTATAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 548
Db 612 ATATTATAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 671
QY 549 TGAAGCGCTTTCATCGGAGCAGTGGGGGATACAAAGCGTGAGAACTGGAGAACTCAT 608
Db 672 TGAAGCGCTTTCATCGGAGCAGTGGGGGATACAAAGCGTGAGAACTGGAGAACTCAT 731
QY 609 CACAGACATGAGCTGGCGCTTCCGAGCTTCTTCAAGGCTTTCAGAGTGTCTGCGAG 668
Db 732 CACAGACATGAGCTGGCGCTTCCGAGCTTCTTCAAGGCTTTCAGAGTGTCTGCGAG 791
QY 669 CTCGAGGGTTCAGGAGATCAAGGACTTCAAGGACTTCTACCTTTCCATAGAGATCA 728
Db 792 CTCGAGGGTTCAGGAGATCAAGGACTTCAAGGACTTCTACCTTTCCATAGAGATCA 851
QY 729 TTAGTAGAAGTTCGGAATGCAAAATACAGTGTGAAGAGAACCTCACCCAGTTATAGG 788
Db 852 TTAGTAGAAGTTCGGAATGCAAAATACAGTGTGAAGAGAACCTCACCCAGTTATAGG 911
QY 789 AGCTATCCGGTTGAGAAATTTG-TGGTACCATGATATCATTTACCTTTGCTTAT 847
Db 912 AGCTATCCGGTTGAGAAATTTG-TGGTACCATGATATCATTTACCTTTGCTTAT 971
QY 848 ATAAGTGAAGACCTGAGAGATGAGCGCCCTGTCAGTACGATATCTGCTTTGATC 907
Db 972 ATAAGTGAAGACCTGAGAGATGAGCGCCCTGTCAGTACGATATCTGCTTTGATC 1031
QY 908 AGAATGACCAAG 918
Db 1032 AGAWKEAAGG 1042

```

```

RESULT 10
AL541167
LOCUS AL541167 Homo sapiens PLACENTA 1201 bp mRNA linear EST 12-MAY-2003
DEFINITION S-PRIME, mRNA sequence.
ACCESSION AL541167
VERSION AL541167.2 GI:30545079
KEYWORDS EST.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 15, 2001 this sequence version replaced gi:12871971.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 131 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 2233.f For
more information about this cluster, see
http://www.genoscope.cns.fr/

```

```

cgi-bin/cluster.cgi?seq=CS0DE005DC06QPI&cluster=2233.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DE005DC06QPI.
FEATURES
location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE005YF12"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-gli3o(dn) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
BASE COUNT 258 a 335 c 313 g 237 t 58 others
ORIGIN
Query Match 76.3%; Score 865.6; DB 9; Length 1201;
Best Local Similarity 98.1%; Pred. No. 1.9e-199;
Matches 895; Conservative 10; Mismatches 4; Indels 3; Gaps 3;
QY 8 CCCAATACGACGCTACAGCTTCGCGAGCTTCCACGGGACGAGCTGATCCGCTCGAGT 67
Db 139 CCCAATACGACGCTACAGCTTCGCGAGCTTCCACGGGACGAGCTGATCCGCTCGAGT 198
QY 68 CGGCTTACC3GCGCGCTG3GCAAGTACAGCGCGGACGACTGGCGGAGAGGCTGCGGCT 127
Db 199 CGGCTTACC3GCGCGCTG3GCAAGTACAGCGCGGACGACTGGCGGAGAGGCTGCGGCT 258
QY 128 ACCTGGAGATCAGCTCGGCTGCGACCGCTTGTGCGGACGAGGAGGCTTTCGCCACC 187
Db 259 ACCTGGAGATCAGCTCGGCTGCGACCGCTTGTGCGGACGAGGAGGCTTTCGCCACC 317
QY 188 GCAACTCAGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCTGCGCGAGCTATCCGAGC 247
Db 318 GCAACTCAGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCTGCGCGAGCTATCCGAGC 377
QY 248 TGGCGCTTTCGGGGGCTGCTGCTGCGCGCGCGCACTGCTTCAAGCGCTGCAAGCAGGCG 307
Db 378 TGGCGCTTTCGGGGGCTGCTGCTGCGCGCGCGCACTGCTTCAAGCGCTGCAAGCAGGCG 437
QY 308 TGGCGAGCTTCCGCGAGTCCCGAGCGCGAGGCTGCTGGCGGAGCTTCCAGCGCGCGG 367
Db 438 TGGCGAGCTTCCGCGAGTCCCGAGCGCGAGGCTGCTGGCGGAGCTTCCAGCGCGCGG 497
QY 368 AGCCTTACAAGTTCCTGAGTTCGCTTACTTCAAGCAAAATATCTCCCAAGCCATCG 427
Db 498 AGCCTTACAAGTTCCTGAGTTCGCTTACTTCAAGCAAAATATCTCCCAAGCCATCG 557
QY 428 CGCTGCTCACACCTTCTACTCAAGCATCCTGATCAGCAATGATGAGAGGAGCAATCG 487
Db 558 CGCTGCTCACACCTTCTACTCAAGCATCCTGATGATGATGATGATGATGATGATGATG 617
QY 488 CATATTATAAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 547
Db 618 CATATTATAAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 677
QY 548 ATGAAGCGCTTTCATCCGAGGAGTGGGGGATACAAACGCGTGAAGAACTGGAGAACATCA 607
Db 678 ATGAAGCGCTTTCATCCGAGGAGTGGGGGATACAAACGCGTGAAGAACTGGAGAACATCA 737
QY 608 TCACAGACATGAGGAGTGGCGCTTCCGAGCTTCTTCAAGCGCTTTTACGAGTGTCTCGCAG 667
Db 738 TCACAGACATGAGGAGTGGCGCTTCCGAGCTTCTTCAAGCGCTTTTACGAGTGTCTCGCAG 797
QY 568 CTTGCGAGGTTTCAGGAGATCAAGGACTTCAAGGATTTCTACCTTTCCATAGCAGATC 727
Db 798 CTTGCGAGGTTTCAGGAGATCAAGGACTTCAAGGATTTCTACCTTTCCATAGCAGATC 857
QY 728 ATTATGTAGAAGTTCGGAATGCAAAATACAGTGTGAAGAGAACCTCACCCAGTTATAG 787

```



```
Db 858 ATTATGTAGAAAGTTCTGGAATGCAAAATACAGTGTGAAGAGAACCTACCCAGTTATAG 917
QY 788 GAGGCTATCCGGTTGAGAAATTTGGGCTACCATATATCATTTGCGAGTTTGCCCTATT 847
Db 918 GAGGCTATCCGGTTGAGAAATTTGGGCTACCATATATCATTTGCGAGTTTGCCCTATT 977
QY 848 ATAAAGTTGAACGACCTGAAGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 907
Db 978 ATAAAGTTGAACGACCTGAAGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1035
QY 908 AGAATGACAAAGG 919
Db 1036 AGAATGACAAAGG 1047

RESULT 11
AL554750
LOCUS
DEFINITION
AL554750 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1086YJ04 5-PRIME, mRNA sequence.
ACCESSION
AL554750
VERSION
AL554750.2 GI:31276560
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 1201)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
JOURNAL
On Feb 15, 2001 this sequence version replaced gi:12895833.
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 2233.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0D1086DE02P1&cluster=2233.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0D1086DE02P1.
Location/Qualifiers
1, 1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1086YJ04"
/tissue_type="PLACENTA"
/tissue="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 272 a 329 c 309 g 250 t 41 others
ORIGIN

Query Match 71.6%; Score 811.6; DB 9; Length 1201;
Best Local Similarity 95.1%; Pred. No. 2.5e-18;
Matches 871; Conservative 18; Mismatches 18; Indels 9; Gaps 5;

QY 8 CCCAATAGAACCTACAGC-TTCCGCGAGCTTCCACGGGACGAGCTGATCCGCTCGAG 66
Db 151 CCCAATAGAACCTACAGC-TTCCGCGAGCTTCCACGGGACGAGCTGATCCGCTCGAG 210
QY 67 TCGGCGCTACCGGACGCGCTGACAAAGTACAGCGCGGAGCACTGG--GCCGAGAGCTGK 124
Db 211 TCGGCGCTACCGGACGCGCTGACAAAGTACAGCGCGGAGCACTGGGCGCGGAGAGCTGG 270
QY 125 GCTACTGTGAGATCAGCTTGGGCTGACCGGCTTCTGCTGCGGACAGCGAGCGCTTCTGCC 184
```

```
Db 271 GCTACTGTGAGATCAGCTTGGGCTGACCGGTTGCTGCGGACACATAKTTCTTCTGCC 330
QY 185 ACCGCAACTGCGAGCGCGCGCGGACCC--GCCGCGGCGCTCGCAGCTATCC 242
Db 331 ACCGCAACTCAKTTCCGCGCGCGCAKCCCKXCCCTGTCCGCGGCGCTCGCAGCTATCC 390
QY 243 CGAGCTGCGGCTTCTTGGGGGCGCTGCTGCGCGGCGGAGCTGCTCAAGCGCTGCAAGCA 302
Db 391 CGAGCTGCGGCTTCTTGGGGGCGCTGCTGCGCGCGGCGGAGCTGCTCAAGCGCTGCAAGCA 450
QY 303 GGGCTTCCGAGCGCTTCCGCGAGTCCCGAGCCCGAGCGGCTGCTGGCGGAGCTTCCAGCG 362
Db 451 GGGCTTCCGAGCGCTTCCGCGAGTCCCGAGTCCCGAGCGGAGCTGCTGGCGGAGCTTCCAGCG 510
QY 363 CCGCGAGCGCTTACAGTTCCTGCGAGTTCCTTCAAGGCAAAATAATCTTCCCAAGCG 422
Db 511 CCGCGAGCGCTTACAGTTCCTGCGAGTTCCTTCAAGGCAAAATAATCTTCCCAAGCG 570
QY 423 CATCGCGGCTGCTCACACCTTTCTACTGAAGCATCTGATGAGCAATCATGAAGAGAA 482
Db 571 CATCGCGGCTGCTCACACCTTTCTACTGAAGCATCTGATGAGCAATCATGAAGAGAA 630
QY 483 CATGGCATATTATAGAGCGCTGCTGCTGCGGAGGACTACATTAAAGAGCTTGGAAACCAA 542
Db 631 CATGGCATATTATAGAGCGCTGCTGCTGCGGAGGACTACATTAAAGAGCTTGGAAACCAA 690
QY 543 GTCATATGAAGCGCTTTCATCGGAGCGTGGGCGGATACCAAGCTGAGAACTGGAGAAC 602
Db 691 GTCATATGAAGCGCTTTCATCGGAGCGTGGGCGGATACCAAGCTGAGAACTGGAGAAC 750
QY 603 ATCCATCAGACATGAGCGTGGCGCTTCCCGACTTCTTCAAGGCGCTTTTACGAGTGTCT 662
Db 751 ATCCATCAGACATGAGCGTGGCGCTTCCCGACTTCTTCAAGGCGCTTTTACGAGTGTCT 810
QY 663 CGAGCGCTGCGAGGGTTCAGGAGATCAAGAGCTTCAAGAGATTTCTTCCATAGC 722
Db 811 CGAGCGCTGCGAGGGTTCAGGAGATCAAGAGCTTCAAGAGATTTCTTCCATAGC 870
QY 723 AGATCATTATGTAAGAGTTCTGGAATGCAAAATACAGTGTGAAGAGAACTCACCCAGT 782
Db 871 AGATCATTATGTAAGAGTTCTGGAATGCAAAATACAGTGTGAAGAGAACTCACCCAGT 930
QY 783 TATAGAGGCTATPCCGGTTGAGAAATTTGCTGCTACCATGTATCATTTACGAGTTGC 842
Db 931 TATAGAGGCTATPCCGGTTGAGAAATTTGCTGCTACCATGTATCATTTACGAGTTGC 990
QY 843 CTATTATAGTTGAACGACCTGAGAAATGCAAGCGGCTGAGAGAGAGAGAGAGAGAGAG 902
Db 991 CTATTATAGTTGAACGACCTGAGAAATGCAAGCGGCTGAGAGAGAGAGAGAGAGAGAG 1046
QY 903 TGATCAGATGACAAG 918
Db 1047 TGATCAGATGACAAG 1062

RESULT 12
AL543090
LOCUS
DEFINITION
AL543090 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1002YP17 5-PRIME, mRNA sequence.
ACCESSION
AL543090
VERSION
AL543090.2 GI:31264937
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 1201)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
JOURNAL
On Feb 15, 2001 this sequence version replaced gi:12875568.
COMMENT
```

Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 2233.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0D1002CH09QP1&cluster=2233.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0D1002CH09QP1.
Location/Qualifiers
1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1002YP17"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 272 a 333 c 326 g 237 t 33 others
ORIGIN
Query Match 71.2%; Score 807.8; DB 9; Length 1201;
Best Local Similarity 98.8%; Pred. No. 2.1e-185;
Matches 839; Conservative 5; Mismatches 2; Indels 3; Gaps 3;
QY 8 CCCAATACGAACGCTACAGCTTCGGAGCTTCCACGGGACGAGCTGATCGCGCTCGAGT 67
DB 138 CCCAATACGAACGCTACAGCTTCGGAGCTTCCACGGGACGAGCTGATCGCGCTCGAGT 197
QY 68 CGGCTACCGGACGCGCTCGAAGTACAGCGGCGAGCACTGGGCGGAGAGCTGCGT 127
DB 198 CGGCTACCGGACGCGCTCGAAGTACAGCGGCGAGCACTGGGCGGAGAGCTGCGT 257
QY 128 ACCTGGAGATCAGCTTCGCGCTGACGAGTTCGCGGAGAGCTGCGGAGAGCTGCGT 187
DB 258 ACCTGGAGATCAGCTTCGCGCTGACGAGTTCGCGGAGAGCTGCGGAGAGCTGCGT 317
QY 188 GCAACTGACGCGCGCGCGAGCGCGGAGCGCGGAGCGCGGAGCTGCGGAGAGCTGCGT 247
DB 318 GCAACTGACGCGCGCGCGAGCGCGGAGCGCGGAGCGCGGAGCTGCGGAGAGCTGCGT 377
QY 248 TGGGCTTTCGGGGGCTGCTGGGCGCGCGGAGCGCGGAGCGCGGAGCTGCGGAGAGCTGCGT 307
DB 378 TGGGCTTTCGGGGGCTGCTGGGCGCGCGGAGCGCGGAGCGCGGAGCTGCGGAGAGCTGCGT 437
QY 308 TGGGCTTTCGGGCTGCGGAGCGCGGAGCGCGGAGCGCGGAGCTGCGGAGAGCTGCGT 367
DB 438 TGGGCTTTCGGGCTGCGGAGCGCGGAGCGCGGAGCGCGGAGCTGCGGAGAGCTGCGT 497
QY 368 AGGCTTACAGTTCCTGAGTTCCTTACCTTACGAGCAATATCTCCCAAGGAGCTGCGT 427
DB 498 AGGCTTACAGTTCCTGAGTTCCTTACCTTACGAGCAATATCTCCCAAGGAGCTGCGT 557
QY 428 CCGCTGCTCACCTTTCCTGAGTTCCTGAGTTCCTGAGTTCCTGAGTTCCTGAGTTCCTGAGT 487
DB 558 CCGCTGCTCACCTTTCCTGAGTTCCTGAGTTCCTGAGTTCCTGAGTTCCTGAGTTCCTGAGT 617
QY 488 CATATTATAGAGCTTCCTGAGTTCCTGAGTTCCTGAGTTCCTGAGTTCCTGAGTTCCTGAGT 547
DB 618 CATATTATAGAGCTTCCTGAGTTCCTGAGTTCCTGAGTTCCTGAGTTCCTGAGTTCCTGAGT 677
QY 548 ATGAAGCTTTCATCCGAGCTGCGGAGCTACACGCTGAGTTCGAGTTCGAGTTCGAGTTCGAGT 607
DB 678 ATGAAGCTTTCATCCGAGCTGCGGAGCTACACGCTGAGTTCGAGTTCGAGTTCGAGTTCGAGT 737
QY 608 TCACAGACATGGAGCTGCGGCTTCCCGAGCTTCTTCAAAGCTTCTTCAAAGCTTCTTCAAAGCT 667

Db 738 TCACAGACATGGAGCTGCGGCTTCCCGAGCTTCTTCAAAGCTTCTTCAAAGCTTCTTCAAAGCT 797
QY 668 CCTCGAGGCTTCCAGGAGATCAAGGAGCTTCAAGGATTTCTACCTTCCATAGCAGATC 727
Db 798 CCTCGAGGCTTCCAGGAGATCAAGGAGCTTCAAGGATTTCTACCTTCCATAGCAGATC 857
QY 728 ATTATGAGAGTTCGGAATGCAAAATACAGTGTGAGAGAACTTCAACCCAGTTATAG 787
Db 858 ATTATGAGAGTTCGGAATGCAAAATACAGTGTGAGAGAACTTCAACCCAGTTATAG 917
QY 788 GAGGCTATCCGTTGAGAAATTTGTGGCTACCATGTATCATTTACTTGCAGTTTGCCTATT 847
Db 918 G-ROCTATCCGTTGAG-ATTTGTGGCTACCATGTATCATTTACTTGCAGTTTG-CTATT 974
QY 848 ATAAGTTGA 856
Db 975 ATAAGTTGA 983

RESULT 13
BU838140
LOCUS
DEFINITION AGENCOURT 8071749 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6088309
5', mRNA sequence.
ACCESSION BU838140
VERSION BU838140.1 GI:24022535
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (Bases 1 to 846)
NIH-MGC http://mgi.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM2326 row: e column: 14
High quality sequence stop: 664.
FEATURES
Location/Qualifiers
1. .846
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6088309"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_112"
/note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dr priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
BASE COUNT 199 a 255 c 215 g 177 t
ORIGIN
Query Match 70.6%; Score 801; DB 13; Length 846;
Best Local Similarity 100.0%; Pred. No. 7.8e-184;
Matches 801; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 126 CTACTGAGATACGCTGCGGTGACCGCTTCTCGCGGACGAGCGAGGCTTCTGCCA 185
Db 1 CTACTGAGATACGCTGCGGTGACCGCTTCTCGCGGACGAGCGAGGCTTCTGCCA 60

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2004, 05:50:18 ; Search time 294.825 Seconds
(without alignments)
10382.969 Million cell updates/sec

Title: US-09-729-674-1_COPY_132_1265
Perfect score: 1134
Sequence: 1 99CGCGCCCAACGAACG.....TGGAACTGAGGAGACCAAC 1134

Scoring table:
IDENTITY_NUC
Gapex 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*
25: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1133.2	99.9	3871	20	AA60801 Human secreted pro
2	1133.2	99.9	3871	22	AA59207 Human cDNA encodin
3	1133.2	99.9	3871	24	AA90876 Human polynucleoti
4	1128.4	99.5	1961	22	AAH14525 Human cDNA sequenc
5	705	62.2	1586	21	AAH31138 Human colon cancer
6	587.6	51.8	818	22	AAH06808 Human cDNA clone (
7	455.4	40.2	2347	24	AB167662 Oesophagus cancer
8	406.4	35.8	408	21	AAH30585 Human colon cancer

9	156.4	13.8	1309	24	AB211344 Human polynucleoti
10	156.4	13.8	2152	24	AB211345 Human polynucleoti
11	156.4	13.8	2211	24	AA517573 DNA encoding novel
12	156.4	13.8	2242	24	AB211345 Human polynucleoti
13	156.4	13.8	2524	22	AAH14513 Human cDNA sequenc
14	156.4	13.8	2563	22	AA93800 Human cDNA encodin
15	156.4	13.8	2585	22	AAH14589 Human cDNA sequenc
16	156.4	13.8	2993	22	AAH16288 Human cDNA sequenc
17	154.8	13.7	2600	21	AAC64725 Human tumour suppr
18	154.8	13.7	2829	21	AAC64724 Human tumour suppr
19	152.6	13.5	2577	22	AA08489 Human secreted pro
20	150.4	13.3	2615	22	AA08507 Human secreted pro
21	147.4	13.0	2322	21	AAC64726 Mouse tumour suppr
22	147.4	13.0	2416	21	AAC64726 Mouse tumour suppr
23	130.8	11.5	2127	24	AA517572 DNA encoding novel
24	126	11.1	618	21	AAZ80131 Human colon cancer
25	121.4	10.7	325	22	AA182020 Human polynucleoti
26	116.2	10.2	1026	24	ABQ26350 Oligonucleotide fo
27	116.2	10.2	1026	24	ABQ26351 Oligonucleotide fo
28	115.8	10.2	1025	24	ABQ27312 Oligonucleotide fo
29	115.8	10.2	1025	24	ABQ27313 Oligonucleotide fo
30	109.2	9.6	2665	25	ACA03965 cDNA downregulated
31	99.2	8.7	652	22	AAF93998 Primer specific fo
32	99.2	8.7	738	22	AAH06612 Human cDNA clone (
33	99	8.7	703	22	AAH06792 Human cDNA clone (
34	98.6	8.7	774	22	AAH06885 Human cDNA clone (
35	98.6	8.7	800	22	AAH06448 Human cDNA clone (
36	98.6	8.7	2789	22	AAH14459 Human cDNA sequenc
37	97.8	8.6	2753	24	ABK94910 Human novel polynu
38	97.8	8.6	2753	24	ABK94933 Lung cancer associ
39	87	7.7	349	21	AAF18376 Oligonucleotide fo
40	81.6	7.2	1025	24	ABQ27310 Oligonucleotide fo
41	81.6	7.2	1025	24	ABQ27311 Oligonucleotide fo
42	81.6	7.2	1026	24	ABQ26352 Oligonucleotide fo
43	81.6	7.2	1026	24	ABQ26353 Oligonucleotide fo
44	73.2	6.5	550	25	ACA57060 Human adipocyte Se
45	65.6	5.8	681	23	AA580583 DNA encoding novel

ALIGNMENTS

RESULT 1

AA60801

ID AA60801 standard; DNA; 3871 bp.

XX AA60801;

XX AA60801;

XX 09-AUG-1999 (first entry)

XX Human secreted protein encoding DNA (clone bd306-7).

XX Secreted protein; kidney; lung; brain; blood; testis; bone marrow;
XX nutritional activity; cytokine; cell proliferation; immune stimulation;
XX hematopoiesis regulation; tissue growth; thrombolytic; gene therapy;
XX anti-inflammatory; tumour invasion; ss.

OS Homo sapiens.

XX Homo sapiens.

XX WO9926961-A1.

XX 03-JUN-1999.

XX 24-NOV-1998; 98WO-US25149.

XX 23-NOV-1998; 98US-019786.

XX 26-NOV-1997; 97US-0066804.

XX (GEMY) GENETICS INST INC.

XX Agostino MJ, Clark HF, Collins-Racie LA, Evans C;

XX Fechtel K, Jacobs K, Lavallie ER, McCoy JM, Merberg D;

XX Steininger RU, Treacy M, Wong GG;

XX WPI: 1999-357809/30.
 DR P-PSDB; AAY17219.
 XX
 PT New polynucleotides encoding secreted proteins
 XX
 XX Claim 1; Page 110-111; 133pp; English.
 XX
 CC The invention relates to secreted proteins (AAY17219-228) encoded by
 CC polynucleotides obtained from human fetal kidney, adult lung, adult
 CC kidney, adult brain, adult blood, adult testes, and fetal brain and
 CC murine adult bone marrow cDNA libraries. The secreted protein nucleic
 CC acid sequences (X6801-811) correspond to clones bd306-7, g1283-6,
 CC fk17-3, x213-2x, na316-1, nf93-20, np164-1, pe204-1, ya13-1 and yb-1.
 CC (all clones are deposited as ATCC 98599); The PNs and proteins are
 CC predicted to have biological activities which would make them suitable
 CC for treating, preventing or ameliorating medical conditions in humans and
 CC animals, although no supporting data is given. Suggested activities
 CC include nutritional activity, cytokine and cell
 CC proliferation/differentiation activity, immune stimulating (e.g. as
 CC vaccines) or suppressing activity, hematopoiesis regulating activity,
 CC tissue growth activity, activin/ inhibin activity,
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,
 CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumour
 CC invasion suppressor activity, and tumour inhibition activity. The PNs are
 CC also stated to be useful for gene therapy.
 XX
 SQ Sequence 3871 BP; 965 A; 913 C; 1033 G; 958 T; 2 other;
 Query Match 99.9%; Score 1133.2; DB 20; Length 3871;
 Best Local Similarity 100.0%; Pred. No. 98-244;
 Matches 1134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 GGGCGCGCCCAATACGAACTTACAGCTTCGCGAGCTTCCACGGGACGAGCTGATCCCG 60
 132 GGGCGCGCCCAATACGAACTTACAGCTTCGCGAGCTTCCACGGGACGAGCTGATCCCG 191
 61 CTCGAGTCGGCTTACCGGACACGGCTGAGCAAGTACAGCGCGGAGCACTGGCGGAGAGC 120
 192 CTCGAGTCGGCTTACCGGACACGGCTGAGCAAGTACAGCGCGGAGCACTGGCGGAGAGC 251
 121 STKGCTACCTGGAGATCAGCTCGGCTGCGCTGACCGCTGCTGCGGACGAGGCTTC 180
 252 STKGCTACCTGGAGATCAGCTCGGCTGCGCTGACCGCTGCTGCGGACGAGGCTTC 311
 181 TGCCACCGCAACTGCGCGCGCGCGCGAGCCCGCGCGCTCGCGAGCTAT 240
 312 TGCCACCGCAACTGCGCGCGCGCGCGAGCCCGCGCGCTCGCGAGCTAT 371
 241 CCGGAGTCGGCTTACCGGCGCGCTGCTGCGCGCGGAGCACTGCTCAAGCGCTGCAAG 300
 372 CCGGAGTCGGCTTACCGGCGCGCTGCTGCGCGCGGAGCACTGCTCAAGCGCTGCAAG 431
 301 CAGGCGCTTCCGAGCTTCCGCGAGTCCCGAGCCGCGGAGGCTGCTGGCGGACTTCCAG 360
 432 CAGGCGCTTCCGAGCTTCCGCGAGTCCCGAGCCGCGGAGGCTGCTGGCGGACTTCCAG 491
 361 CCGCGGAGCCCTACAGTTCCTGAGTTCCTGAGTTCCTGAGTTCCTGAGTTCCTGAGTTC 420
 492 CCGCGGAGCCCTACAGTTCCTGAGTTCCTGAGTTCCTGAGTTCCTGAGTTCCTGAGTTC 551
 421 GCCATCGCGCTGCTCACACCTTCTACTGAGGATCTGATGACGAAATGATGAGAGG 480
 552 GCCATCGCGCTGCTCACACCTTCTACTGAGGATCTGATGACGAAATGATGAGAGG 611
 481 AACATGGCATATTATAGAGCTGCTGGTGGCGGAGCTACATTAAGACCTGGAAACC 540
 612 AACATGGCATATTATAGAGCTGCTGGTGGCGGAGCTACATTAAGACCTGGAAACC 671
 541 AAGTCATATGAAGCTTCTTATCCGAGCTGCGGATACACCGTGTGAGACTGGAGA 600
 672 AAGTCATATGAAGCTTCTTATCCGAGCTGCGGATACACCGTGTGAGACTGGAGA 731

QY 601 ACATCCATCACAGACATGGAGCTGGCCCTTCCCGACTTCTTCAAAGCCCTTTTACAGTGT 660
 DB 732 ACATCCATCACAGACATGGAGCTGGCCCTTCCCGACTTCTTCAAAGCCCTTTTACAGTGT 791
 QY 661 CTCGAGCGCTGCGAGGTTCCAGGGAGATCAAGGACTTCAAGGATTTCTTACCTTTCCATA 720
 DB 792 CTCGAGCGCTGCGAGGTTCCAGGGAGATCAAGGACTTCAAGGATTTCTTACCTTTCCATA 851
 QY 721 GCAGATCATTTATGTAGAGTTCTTGAATGCAAAATACAGTGTGAAGAGAACTCACCCCA 780
 DB 852 GCAGATCATTTATGTAGAGTTCTTGAATGCAAAATACAGTGTGAAGAGAACTCACCCCA 911
 QY 781 GTTATAGAGGCTATCCCGTTGAGAAATTTGTGGCTACCATGTATCATTTCTGAGTTT 840
 DB 912 GTTATAGAGGCTATCCCGTTGAGAAATTTGTGGCTACCATGTATCATTTCTGAGTTT 971
 QY 841 GCCTATTATAGTTGAACGACCTGAAGAAATCAGCCCCCTGTGAGTCAGTATCTGCTC 900
 DB 972 GCCTATTATAGTTGAACGACCTGAAGAAATCAGCCCCCTGTGAGTCAGTATCTGCTC 1031
 QY 901 TTTGATCAGATGACAGGTTCTATGCGAGAACCTGGTGTATTACCATGACGAGGAC 960
 DB 1032 TTTGATCAGATGACAGGTTCTATGCGAGAACCTGGTGTATTACCATGACGAGGAC 1091
 QY 961 ACTTGGGCGCTTCCGATGAGCACTTCCAGCCCGAGCTTGAAGCAGTTTCTTTTAA 1020
 DB 1092 ACTTGGGCGCTTCCGATGAGCACTTCCAGCCCGAGCTTGAAGCAGTTTCTTTTAA 1151
 QY 1021 GTGACACACTCCAGAGGAGCTGTATGACTTTGCTAAGGAAATATATATGATGATGAT 1080
 DB 1152 GTGACACACTCCAGAGGAGCTGTATGACTTTGCTAAGGAAATATATATGATGATGAT 1211
 QY 1081 GAGGAGAGAGTTGTGGAATATGTGGATGACCTCTTGGAACTGGAGGAGACGAGC 1134
 DB 1212 GAGGAGAGAGTTGTGGAATATGTGGATGACCTCTTGGAACTGGAGGAGACGAGC 1265
 RESULT 2
 AAS59207
 ID AAS59207 standard; cDNA; 3871 BP.
 XX
 AC AAS59207;
 XX
 XX 16-JAN-2002 (first entry)
 XX
 DE Human cDNA encoding a secreted protein bd306_7.
 XX
 KW Human; secreted protein; ss; antiinflammatory; immunosuppressive;
 KW neutrotropic; neuroprotective; antiarthritic; antimicrobial; vulnary;
 KW cytostatic; antidiabetic; virucide; antifertility; anticonvulsant;
 KW vasotropic; antiparkinsonian; immunostimulant; dermatological;
 KW antirheumatic; antitumor; antitumor; osteopathic; tranquilizer;
 KW cerabroprotective; cytokine; cell proliferation; cell differentiation;
 KW immune deficiency; severe combined immunodeficiency; SCID; tumour;
 KW autoimmune disorder; multiple sclerosis; rheumatoid arthritis;
 KW graft-versus-host disease; myeloid deficiency; wound healing; ulcer;
 KW periodontal disease; osteoporosis; osteoarthritis; Alzheimer's disease;
 KW Parkinson's disease; Huntington's disease; infection; cardiac disease;
 KW stroke; sepsis; inflammatory bowel disease; contraceptive; immunogen;
 XX
 OS Homo sapiens.
 XX
 XX WO200175068-A2.
 XX
 PD 11-OCT-2001.
 XX
 XX 22-MAR-2001; 2001WO-US09369.
 XX
 XX 30-MAR-2000; 2000US-0539330.
 PR
 PR 04-DEC-2000; 2000US-0729674.
 XX
 XX (GEMY) GENETICS INST INC.
 PA

PD 08-NOV-2001.
 XX 04-DEC-2000; 2000US-0729674.
 XX 10-APR-1997; 97US-126425P.
 PR 04-DEC-1997; 97US-067454P.
 PR 20-DEC-1997; 97US-068379P.
 PR 02-JAN-1998; 98US-070346P.
 PR 07-JAN-1998; 98US-070643P.
 PR 08-JAN-1998; 98US-070755P.
 PR 13-JAN-1998; 98US-071304P.
 PR 22-JAN-1998; 98US-072113P.
 PR 30-JAN-1998; 98US-073095P.
 PR 18-FEB-1998; 98US-075038P.
 PR 30-MAR-2000; 2000US-0539330.
 PR 23-NOV-1998; 98US-0197886.
 XX (JACO/) JACOBS K.
 PA (MCCO/) MCCOY J M.
 PA (LAVA/) LAVALLIE E R.
 PA (COLL/) COLLINS-RACIE L A.
 PA (EVAN/) EVANS C.
 PA (MERB/) MERBERG D.
 PA (TREA/) TREACY M.
 PA (AGOS/) AGOSTINO M J.
 PA (STEI/) STEININGER R J.
 PA (SPAU/) SPAULDING V.
 PA (WONG/) WONG G G.
 PA (CLAR/) CLARK H.
 PA (FECH/) FECHTEL K.
 XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C,
 PI Merberg D, Treacy M, Agostino MJ, Steininger RJ, Spaulding V;
 PI Wong GG, Clark H, Fechtel K;
 XX WPI: 2002-040725/05.
 DR P-PSDB: ABB55698.
 DR
 XX New secreted proteins and encoding polynucleotides, useful in gene
 PT therapies, particularly for preventing or treating autoimmune
 PT disorders, cancer, graft-versus-host disease, wound, osteoporosis,
 PT stroke or inflammations
 XX Claim 1: Page 175-177; 349pp; English.
 XX The invention relates to isolated polynucleotides (ABA90876-ABA90968 and
 CC ABA90980) and encoded proteins (ABB55698-ABB55800), especially
 CC polynucleotides SEQ ID NO 1 (ABA90876) and SEQ ID NO 19 (ABA90885) and
 CC proteins SEQ ID NO 2 (ABB55698) and SEQ ID NO 20 (ABB55707) contained in
 CC clones bd306-7 and Yb8-1 respectively and the clones bd306-7 and Yb8-1
 CC are deposited with the American Type Culture Collection (ATCC) with
 CC accession number 98599. The polynucleotides and encoded polypeptides have
 CC cytostatic, anti-inflammatory, immunomodulatory, vulnerary,
 CC neuroprotective, activin, inhibin, chemotactic, haemostatic, thrombolytic
 CC and anti-inflammatory activity and acting as cytokine modulators,
 CC haematopoiesis regulators, tissue growth modulators and/or cadherin
 CC suppressors. The polypeptides and polynucleotides are useful in gene
 CC therapies, particularly for preventing, treating or ameliorating any of
 CC the following diseases: immune deficiency and disorders; e.g. bacterial
 CC or fungal infections, autoimmune disorders, cancer, systemic lupus
 CC erythematosus or graft-versus-host disease; myeloid or lymphoid cell
 CC deficiencies; wound, burns, incisions and ulcers, osteoporosis or
 CC osteoarthritis; central and peripheral nervous system diseases and
 CC neuropathies; e.g. Alzheimer's, Parkinson's disease, Huntington's
 CC disease, amyotrophic lateral sclerosis or Shy-Drager syndrome;
 CC haemophilia, cardiac infarction or stroke; inflammations, shock, sepsis
 CC or systemic inflammatory response syndrome, ischaemia-reperfusion
 CC injury, endotoxin lethality, arthritis, inflammatory bowel disease or
 CC Crohn's disease; or tumours or cancers, pemphigus vulgaris or pemphigus
 CC foliaceus.
 XX Sequence 3871 BP; 965 A; 913 C; 1033 G; 958 T; 2 other;

Query Match 99.9%; Score 1133.2; DB 24; Length 3871;
 Best Local Similarity 100.0%; Pred. No. 9e-244;
 Matches 1134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGCGCCCCCAATACGAAACGCTACAGCTTCGGCAGCTTCCACGGGACGAGCTGATGCGG 60
 DB 132 GGGCGCCCCCAATACGAAACGCTACAGCTTCGGCAGCTTCCACGGGACGAGCTGATGCGG 191
 QY 61 CTCGAGTCGGCTTACCGGCAACGCTTGGCAAGTACAGCGCGGAGCACTGGGCGGAGAGC 120
 DB 192 CTCGAGTCGGCTTACCGGCAACGCTTGGCAAGTACAGCGCGGAGCACTGGGCGGAGAGC 251
 QY 121 STKGCTACCTGAGATACAGCTTGGGCTGCAACGCTTGTGCGGAGCAGCGAGGCTTC 180
 DB 252 STKGCTACCTGAGATACAGCTTGGGCTGCAACGCTTGTGCGGAGCAGCGAGGCTTC 311
 QY 181 TGCCACCGCAACTGCAGCGCGCGCCAGCGCGAGCGCGCGCGCGCGCTTCGCGAGCTAT 240
 DB 312 TGCCACCGCAACTGCAGCGCGCGCCAGCGCGAGCGCGCGCGCGCGCTTCGCGAGCTAT 371
 QY 241 CCCGAGCTGGCGCTTCTTGGGGGGCTTCTGCGCGCGCGCGCGCGCGCTCAAGCGCTGCAAG 300
 DB 372 CCCGAGCTGGCGCTTCTTGGGGGGCTTCTGCGCGCGCGCGCGCGCTCAAGCGCTGCAAG 431
 QY 301 CAGGCGCTGCCAGCGCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTCGCGAG 360
 DB 432 CAGGCGCTGCCAGCGCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTCGCGAG 491
 QY 361 CGCGCGAGCGCTTACAAAGTCTTCTGCGAGTTCCTTCAAGGCAATAATCTCCCCAAA 420
 DB 492 CGCGCGAGCGCTTACAAAGTCTTCTGCGAGTTCCTTCAAGGCAATAATCTCCCCAAA 551
 QY 421 GCCATCGCGCTGCTCACACCTTCTTACTGAAGCATCTCTGATGACGAAATGATGAAGAGG 480
 DB 552 GCCATCGCGCTGCTCACACCTTCTTACTGAAGCATCTCTGATGACGAAATGATGAAGAGG 611
 QY 481 AACATGCGCATATTATTAAGAGCGCTTCTGCGCGCGCGCGCGCGCGCTTCAAGACCTTGAAGAAC 540
 DB 612 AACATGCGCATATTATTAAGAGCGCTTCTGCGCGCGCGCGCGCGCTTCAAGACCTTGAAGAAC 671
 QY 541 AAGTCATATGAAGCGCTTCTTCAAGAGAGTGGGGGCGATACACGCTGAGAACTGAGA 600
 DB 672 AAGTCATATGAAGCGCTTCTTCAAGAGAGTGGGGGCGATACACGCTGAGAACTGAGA 731
 QY 601 ACATCGATCACAGATGAGCTGCGCTTCCGCGCTTCTTCAAGACCTTCTTACGAGTGT 660
 DB 732 ACATCGATCACAGATGAGCTGCGCTTCCGCGCTTCTTCAAGACCTTCTTACGAGTGT 791
 QY 661 CTCGCGAGCTCGAGGGTTCAGGGAGATCAAGGACTTCAAGGATTTCTACCTTTCCATA 720
 DB 792 CTCGCGAGCTCGAGGGTTCAGGGAGATCAAGGACTTCAAGGATTTCTACCTTTCCATA 851
 QY 721 GCAGATCATATTATGAGAGTTCGGAATGCAAAATACAGTGTGAAGAGAACTCACCACCA 780
 DB 852 GCAGATCATATTATGAGAGTTCGGAATGCAAAATACAGTGTGAAGAGAACTCACCACCA 911
 QY 781 GTTATAGAGGCTATCCGGTTGAGAAATTTGTGCTACCATGTATCATTTACGATTT 840
 DB 912 GTTATAGAGGCTATCCGGTTGAGAAATTTGTGCTACCATGTATCATTTACGATTT 971
 QY 841 GCCTATTATAGTTGAACGACATGAGNATGACAGCCCCCTGTGAGTCAGCTATCTGCTC 900
 DB 972 GCCTATTATAGTTGAACGACATGAGNATGACAGCCCCCTGTGAGTCAGCTATCTGCTC 1031
 QY 901 TTTGATCAGATGACAAAGTTCATGACGAGAACTTGTGTATTACCACTACCAAGGAC 960
 DB 1032 TTTGATCAGATGACAAAGTTCATGACGAGAACTTGTGTATTACCACTACCAAGGAC 1091
 QY 961 ACTTGGGGCTTCTCGGATGAGCACTTCCAGCGCGAGACTGAGAGGTCAGTTCTTTAAT 1020
 DB 1092 ACTTGGGGCTTCTCGGATGAGCACTTCCAGCGCGAGACTGAGAGGTCAGTTCTTTAAT 1151
 QY 1021 GTGACCACTCCAGAGGAGCTGTATGACTTTTGCTAAGGAAATATATAATGATGATGAT 1080

Db 1152 GTGACACACATCCAGAGAGCTGTATGACTTTGCTTAAGGAAATATATATGATGATGAT 1211
Qy 1081 GAGGGAGAACTTGTGGAATATGTGGATGACCTCTTGGAACTGGAGGAGACCAAGC 1134
Db 1212 GAGGGAGAACTTGTGGAATATGTGGATGACCTCTTGGAACTGGAGGAGACCAAGC 1265

RESULT 4
AAH14525
ID AAH14525 standard; cDNA; 1961 BP.
XX
AC AAH14525;
XX
DT 25-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:12066.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.

XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 8; SEQ ID 12066; 2537pp + CD ROM; English.

XX
CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

XX
SQ Sequence 1961 BP; 417 A; 579 C; 481 G; 484 T; 0 other;

Query Match 99.5%; Score 1128.4; DB 22; Length 1961;
Best Local Similarity 99.6%; Pred. No. 8.9e-243;
Matches 1129; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGGCGCGCCCAATACGAACCTACAGCTTCGGCAGCTTCCACGGGACGAGCTGATGCCG 60
Db 109 GGGCGCGCCCAATACGAACCTACAGCTTCGGCAGCTTCCACGGGACGAGCTGATGCCG 168

Qy 61 CTCGAGTCGGCCTACCGCGCGCGCTGGCAAGTACAGCGCGGAGCACTGGGCGGAGAGC 120
Db 169 CTCGAGTCGGCCTACCGCGCGCGCTGGCAAGTACAGCGCGGAGCACTGGGCGGAGAGC 228

Qy 121 STKGCTACCTGGAGATCAGCTTCGGCTGCACCGCTTGTGCGCGACGAGGCGCTTC 180
Db 229 GTGGCTACCTGGAGATCAGCTTCGGCTGCACCGCTTGTGCGCGACGAGGCGCTTC 288

Qy 181 TGCCACCGCAACTCAGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCTTCGCGAGCTAT 240
Db 289 TGCCACCGCAACTCAGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCTTCGCGAGCTAT 348

Qy 241 CCGAGCTCGGCTCTTCGGGGGCGCTGCTGCGCGCGCGCGCACTGCTCTCAAGCGCTGCAAG 300
Db 349 CCGAGCTCGGCTCTTCGGGGGCGCTGCTGCGCGCGCGCGCACTGCTCTCAAGCGCTGCAAG 408

Qy 301 CAGGCGCTGCGAGCTTCGGCGAGTCCCGCAGTCCCGCAGCGCGCGCGCGCGCTTCGAG 360
Db 409 CAGGCGCTGCGAGCTTCGGCGAGTCCCGCAGTCCCGCAGCGCGCGCGCGCGCTTCGAG 468

Qy 361 CGCGCGAGCGCTCAAGTTCTCTGAGTTCGCTTACTTCAAGGCAATAATATCTCCCAAA 420
Db 469 CGCGCGAGCGCTCAAGTTCTCTGAGTTCGCTTACTTCAAGGCAATAATATCTCCCAAA 528

Qy 421 GCGATCGCGCTGCTCAACCTTCTACTGAGCATCCTGATGACGAAATGATGAGAGG 480
Db 529 GCGATCGCGCTGCTCAACCTTCTACTGAGCATCCTGATGACGAAATGATGAGAGG 588

Qy 481 AACATGGCATATTATAAGAGCGCTCGCTGGTCCCGAGGACTACATTAAGACCTTGGAAAC 540
Db 589 AACATGGCATATTATAAGAGCGCTCGCTGGTCCCGAGGACTACATTAAGACCTTGGAAAC 648

Qy 541 AAGTCATATGAAGCGCTGTTTCATCCGAGCAGTGGGGCGCATACACGGTGAGAACTGAGA 600
Db 649 AAGTCATATGAAGCGCTGTTTCATCCGAGCAGTGGGGCGCATACACGGTGAGAACTGAGA 708

Qy 601 ACATCCATACAGACATGGAGCTGGCGCTTCCCGACTTCTTCAAGGCTTTTACGAGTGT 660
Db 709 ACATCCATACAGACATGGAGCTGGCGCTTCCCGACTTCTTCAAGGCTTTTACGAGTGT 768

Qy 661 CTCGAGCGCTCGAGGGTTCCAGGGAGATCAAGGACTTCAAGGATTTCTACCTTTCCATA 720
Db 769 CTCGAGCGCTCGAGGGTTCCAGGGAGATCAAGGACTTCAAGGATTTCTACCTTTCCATA 828

Qy 721 GCAGATCATATGTAGAGTTCTGGAATGCAAAATACAGTGTGAAGAGAACCTCACCCCA 780
Db 829 GCAGATCATATGTAGAGTTCTGGAATGCAAAATACAGTGTGAAGAGAACCTCACCCCA 888

Qy 781 GTTATAGAGGCTATCCGGTTGAGAAATTTGTGGCTACCATGTATCATTTACTTGCAGTTT 840
Db 889 GTTATAGAGGCTATCCGGTTGAGAAATTTGTGGCTACCATGTATCATTTACTTGCAGTTT 948

Qy 841 GCCTATTATAGTTGAACGACCTGAAGAAATGACGCCCTGTGCGAGTCAGCTATCTGCTC 900
Db 949 GCCTATTATAGTTGAACGACCTGAAGAAATGACGCCCTGTGCGAGTCAGCTATCTGCTC 1008

Qy 901 TTTGATCAGATGACAAAGTTCATCAGCAGAACCTGCTGTATTATACCAAGTACCACAGGGAC 960
Db 1009 TTTGATCAGATGACAAAGTTCATCAGCAGAACCTGCTGTATTATACCAAGTACCACAGGGAC 1068

Qy 961 ACTTGGGCGCTCTCGATGAGCACTTCCAGCCAGACCTGAAGCAGTTTCACTTTTAAAT 1020
Db 1069 ACGTGGGCGCTCTCAGATGAGCACTTCCAGCCAGACCTGAAGCAGTTTCACTTTTAAAT 1128

Qy 1021 GTGACCACTCCAGAGGAGCTGTATGACTTTTCTTAAGGAAATATAATGATGATGAT 1080

Db	1129	GTGACCACACTCCAGAGGAGCTGTATGACTTTTCTAAGGAAATATAATGGATGATGAT	1188
Qy	1081	GAGGGAGAAGTTGGGAATATGTGGATGACCTCTTGGAACTGGGAGGAGACGACG	1134
Db	1189	GAGGGAGAAGTTGGGAATATGTGGATGACCTCTTGGAACTGGGAGGAGACGACG	1242
RESULT 5			
AAH31138			
ID	AAH31138	standard; cDNA; 1586 BP.	
XX	AAH31138;		
AC	AAH31138;		
XX	27-JUL-2001	(first entry)	
DT			
XX	Human colon cancer cell line Km12L4-A	cDNA library derived seq#1072.	
XX	Human; diagnosis; colon cancer; cancer; malignant; chromosome mapping;		
KW	detection; colon cancer cell line Km12L4-A; ss.		
XX	Homo sapiens.		
OS			
XX	WO200018916-A2.		
PN	06-APR-2000.		
XX	23-SEP-1999;	39WO-US22226.	
PF	28-SEP-1998;	98US-0102161.	
XX	28-SEP-1998;	98US-0102180.	
PR	29-SEP-1998;	98US-0102380.	
PR	08-OCT-1998;	98US-0103815.	
PR	27-OCT-1998;	98US-0105877.	
XX	(CHIR) CHIRON CORP.		
PA	(HYSE-) HYSEQ INC.		
XX	Williams LT, Escobedo J, Inniss MA, Garcia PD, Sudduth-Klinger J;		
PI	Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;		
PI	Lamson G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;		
PI	Leshkowitz D, Kita D, Garcia V, Jones LW, Stache-Crain B;		
XX	WPI; 2000-293155/25.		
DR			
XX	Polynucleotide library comprising 1079 defined sequences, useful in		
PT	the form of an array to detect cancer or susceptibility to cancer -		
XX	Claim 1; Page 498-499; 502pp; English.		
FS			
XX	The present invention describes a library of polynucleotides comprising		
CC	1079 nucleotide sequences (given in AAH30067 to AAH31145). Also described		
CC	are: (1) an isolated polynucleotide (I) having at least 90% identity to		
CC	one of the 1079 sequences; (2) a recombinant host cell containing (I);		
CC	(3) an isolated polypeptide (II) encoded by (I); (4) an antibody that		
CC	specifically binds to (II); (5) a vector comprising (I); and (6) a method		
CC	of detecting differentially expressed genes correlated with a cancerous		
CC	state of a mammalian cell comprising detecting a gene product encoded by		
CC	65 of the 1079 sequences given in the specification. The polynucleotides		
CC	are used to monitor patients having (or susceptible) to cancer to detect		
CC	potentially malignant events at a molecular level before they are		
CC	detectable at a gross morphological level. The polynucleotides are also		
CC	useful for monitoring the efficacy of various therapies and preventive		
CC	interventions. Polynucleotide probes based on the disclosed sequences		
CC	are useful for chromosome mapping and detection of transcription levels.		
CC	The 1079 polynucleotide sequences were derived from a human colon cancer		
CC	cell line Km12L4-A cDNA library.		
XX			
XX	Sequence 1586 BP; 381 A; 406 C; 340 G; 453 T; 6 other;		
SQ			
Query Match	62.2%;	Score 705;	DB 21; Length 1586;
Best Local Similarity	99.3%;	Pred. No. 4.4e-148;	
Matches 702; Conservative	5; Mismatches 0;	Indels 0; Gaps	

PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 1; SEQ ID 3643; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB32446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 818 BP; 159 A; 271 C; 236 G; 148 T; 4 other;

Query Match
Best Local Similarity 51.8%; Score 587.6; DB 22; Length 818;
Matches 631; Conservative 2; Mismatches 18; Indels 5; Gaps 3;

QY 1 GGGCGCGCCCAATACGAAGCTACAGCTTCCCGAGCTTCCACGGGACGAGCTGATGCCG 60
DB 109 GGGCGCGCCCAATACGAAGCTACAGCTTCCCGAGCTTCCACGGGACGAGCTGATGCCG 168
QY 61 CTCGAGTCGGCTTACCGGACCGGCTGGCAAGTACAGCGGAGCACTGGGCGCGAGAGC 120
DB 169 CTCGAGTCGGCTTACCGGACCGGCTGGCAAGTACAGCGGAGCACTGGGCGCGAGAGC 228
QY 121 SYKGGCTACTGGAGATCAGCTGCGGTGACCGCTTCCCGAGCTTCCCGAGCGAGCGCTTC 180
DB 229 GTGGGCTACTGGAGATCAGCTGCGGTGACCGCTTCCCGAGCTTCCCGAGCGAGCGCTTC 288
QY 181 TGCACCGCAACTGCGAGCGCGCGCGCGAGCCGAGCGCGCGCGCGCGCGCGCGCTAT 240
DB 289 TGCACCGCAACTGCGAGCGCGCGCGCGAGCCGAGCGCGCGCGCGCGCGCGCTAT 348
QY 241 CCCGAGCTCGGCTCTTCCGGGGCGCTTCCCGCGCGCGCGCGCGCTCAAGCGCTGCAAG 300
DB 349 CCCGAGCTCGGCTCTTCCGGGGCGCTTCCCGCGCGCGCGCGCTCAAGCGCTGCAAG 408
QY 301 CAGGGCTGCGAGCTTCCCGAGTCCAGCCAGCGCGAGGTCTCGCGGACTTCCAG 360
DB 409 CAGGGCTGCGAGCTTCCCGAGTCCAGCCAGCGCGAGGTCTCGCGGACTTCCAG 468
QY 361 CCGCGCGAGCGCTTCAAGTCTCGAGTTCGCTTACTTCAAGGCAATAATCTCCCAAAA 420
DB 469 CCGCGCGAGCGCTTCAAGTCTCGAGTTCGCTTACTTCAAGGCAATAATCTCCCAAAA 528

QY 421 GCCATCGCGCTGCTCACACCTTTTCTACTGAAGCATCCTGATGACGAAATGATGAAGG 480
DB 529 GCCATCGCGCTGCTCACACCTTTTCTACTGAAGCATCCTGATGACGAAATGATGAAGG 588
QY 481 AACATGCGATATTATTAAGAGCGCTGCTGTCGCGAGGACTACATTAAAGACCTTGAAGAC 540
DB 589 AACATGCGATATTATTAAGAGCGCTGCTGTCGCGAGGACTACATTAAAGACCTTGAAGAC 647
QY 541 AAGT--CATATGAAAGCCTGTTTCATCGAGCAGTGGCGGCATACAAACGGTGAGAACTGGA 598
DB 648 AAGTTCATATGAAGAGCCTGTTTCATCGGANCAGTGGCGGCATACAAACGGTGAGAACTGGA 707
QY 599 GAACATCCATCAGACAGATGAGCTGGCC--TTCCGAGCTTCTTCAAAGCCTTTT 652
DB 708 GAACATTCATCAGACAGATGAGCTGGCCCTTTCCGAGCTTCTTCAAAGAGNCTTT 763

RESULT 7
ABL67662
ID ABL67662 standard; DNA; 2347 BP.
XX ABL67662;
XX
DT 15-MAY-2002 (first entry)
XX
DE Oesophagus cancer related gene sequence SEQ ID NO:5999.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
FN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PP 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235113P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.

PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244857P.
 PR 01-NOV-2000; 2000US-245084P.
 XX
 PA (AVAL-) AVALON PHARM.
 XX
 XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 XX
 DR WPI; 2002-188264/24.
 XX
 PT Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set -
 XX
 XX Claim 1; SEQ ID 5999; 44pp; English.
 PS
 CC The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilm's tumour.
 XX
 SQ Sequence 2347 BP; 517 A; 689 C; 733 G; 408 T; 0 other;
 Query Match. 40.2%; Score 455.4; DB 24; Length 2347;
 Best Local Similarity 65.1%; Pred. No. 3.3e-92;
 Matches 723; Conservative 2; Mismatches 368; Indels 18; Gaps 3;
 5 GCGCCCAATACGAACGCTCAGCTTCCGACGTTCCACCGGACGAGCTGATGCCGCTCG 64
 61 GGGCGCAGTACGAGNAGTACAGCTTCCGGGCTTCCCGCCGAGACCTGATGCCGCTGG 120
 65 AGTCGGCTACCGGACGCGCTGGACAAAGTACAGCGCGGAGACTGGCGGAGAGCGTGG 124
 121 CCGGGGGGTACGGGACGCTCTGGAGCAGTACGAGGAGAGAGCTGGCGGAGAGCGCG 180
 125 GCTACCTGGAGATCAGCTCGGCTGACCGCTTGTCTGCGGACGAGAGGCTTCTGGCC 184
 181 GCTACCTGGAGGCGGCTCGGCTGCACCGCTTCTGCGGACGAGAGGCTTCTGGCC 240
 185 ACCGCAACTCGAGCG-----CCGCGCGAGCGCGGAGCGCGCGCGGCTCGCCA 235
 241 ACGCAACTCGAGCGCGCGCGCGCGCGCGCAAGCGCGATCCGACGCGCGCGCGCGAG 300
 236 GCTA-----TCCGAGCTGGCTCTTCGGGGGCTGTGCGCGCGCGCTGCTCA 289
 301 ACGAGTGGGCTGGAGCTGGCTCTTCGGGGGCTGTGCGCGCGCGCTGCTGCTGC 360
 290 AGCGCTGCAAGCGGCTGCGAGCTTCCGCCAGTCCCGAGCGCGCGAGGCTGCTG 349
 361 GCGCTGCAAGCGGACGCTGCGCGCTTCCAGGTGCTTACCGCGCGCGCGAGCTGCTG 420
 350 CGGACTTCCAGCGCGCGGAGCGCTTACAGTCTTCGCGAGTTCGCTTACTTCAAGGCAATA 409

RESULT 8

AAH30585

ID AAH30585 standard; cDNA; 408 BP.

XX AAH30585;

XX AC

XX 27-JUL-2001 (first entry)

XX DE

XX Human colon cancer cell line Kml2L4-A cDNA library derived sequence #519.

XX Human; diagnosis; colon cancer; cancer; malignant; chromosome mapping;

XX KW detection; colon cancer cell line Kml2L4-A; ss.

XX OS Homo sapiens.

XX XX

XX PN W0200018916-A2.

XX PD

XX 06-APR-2000.

XX XX

XX 23-SEP-1999; 99WO-US22226.

XX XX

PR 28-SEP-1998; 98US-0102161.
PR 28-SEP-1998; 98US-0102180.
PR 29-SEP-1998; 98US-0102380.
PR 08-OCT-1998; 98US-0103815.
PR 27-OCT-1998; 98US-0105877.
XX
XX (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
PI Reinhard C, Giese K, Randazzo P, Kennedy GC, Pot D, Kassam A;
PI Lanson G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;
PI Leeshkowitz D, Kita D, Garcia V, Jones LW, Stache-Crain B;
XX
XX WPI; 2000-293155/25.
XX
XX Polynucleotide library comprising 1079 defined sequences, useful in
PT the form of an array to detect cancer or susceptibility to cancer -
XX
XX Claim 1; Page 333; 502pp; English.
XX
XX The present invention describes a library of polynucleotides comprising
CC 1079 nucleotide sequences (given in AAH30067 to AAH31145). Also described
CC are: (1) an isolated polynucleotide (I) having at least 90% identity to
CC one of the 1079 sequences; (2) a recombinant host cell containing (1);
CC (3) an isolated polypeptide (II) encoded by (1); (4) an antibody that
CC specifically binds to (II); (5) a vector comprising (1); and (6) a method
CC of detecting differentially expressed genes correlated with a cancerous
CC state of a mammalian cell comprising detecting a gene product encoded by
CC 65 of the 1079 sequences given in the specification. The polynucleotides
CC are used to monitor patients having (or susceptible) to cancer to detect
CC potentially malignant events at a molecular level before they are
CC detectable at a gross morphological level. The polynucleotides are also
CC useful for monitoring the efficacy of various therapies and preventive
CC interventions. Polynucleotide probes based on the disclosed sequences
CC are useful for chromosome mapping and detection of transcription levels.
CC The 1079 polynucleotide sequences were derived from a human colon cancer
CC cell line Kml12L4-A cDNA library.
XX
SQ Sequence 408 BP; 117 A; 96 C; 98 G; 97 T; 0 other;

Query Match 35.8%; Score 406.4; DB 21; Length 408;
Best Local Similarity 95.8%; Pred. No. 1.9e-81;
Matches 407; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 428 CCGCTGCTCACACCTTCTACTGAGCATCTGATGACGAATGATGAGGAGACATGG 487
Db 1 CCGCTGCTCACACCTTCTACTGAGCATCTGATGACGAATGATGAGGAGACATGG 60
Qy 488 CATATTATAAGAGCTGCTGCTGGTCCGAGGACTACATTAAAGACCTTGGAAACCAAGTCAT 547
Db 61 CATATTATAAGAGCTGCTGCTGGTCCGAGGACTACATTAAAGACCTTGGAAACCAAGTCAT 120
Qy 548 ATGAAGCCCTTTTCATCCGAGCAGTGGCGCATACACGGTGAGACTGAGACATCCA 607
Db 121 ATGAAGCCCTTTTCATCCGAGCAGTGGCGCATACACGGTGAGACTGAGACATCCA 180
Qy 608 TCACAGACATGGAGCTGGCCCTTCCCGACCTTCTTCAAAGCCTTTTACGAGTGTCTGCGAG 667
Db 181 TCACAGACATGGAGCTGGCCCTTCCCGACCTTCTTCAAAGCCTTTTACGAGTGTCTGCGAG 240
Qy 668 CCTCGAGGGTTTCAGGAGATCAAGGACTTCAAGGATTTCTACCTTTCCATGACGATC 727
Db 241 CCTCGAGGGTTTCAGGAGATCAAGGACTTCAAGGATTTCTACCTTTCCATGACGATC 300
Qy 728 ATTATGTAGAGTTCTGGAATGCAAAATACAGTGTGAAGAGACCTCACCCAGTTATAG 787
Db 301 ATTATGTAGAGTTCTGGAATGCAAAATACAGTGTGAAGAGACCTCACCCAGTTATAG 360
Qy 788 GAGGCTATCCGGTTGAGAAATTTGGCTACCATGTATCTATTCTTC 835
Db 361 GAGGCTATCCGGTTGAGAAATTTGGCTACCATGTATCTATTCTTC 408

RESULT 9
AB211344
ID AB211344 standard; cDNA; 1309 BP.
XX
XX AC AB211344;
XX
XX DT 20-JAN-2003 (first entry)
XX
XX Human polynucleotide SEQ ID NO 226.
XX
XX Human: genome mapping; gene therapy; food supplement; virus; fungus;
KW cell-proliferative disorder; neurodegenerative disease; bacterial;
KW Parkinson's disease; Alzheimer's disease; autoimmune disease;
KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
KW arthritis; cystostatic; immunomodulator; neotropic; neuroprotective;
KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
KW haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
KW antiarthritic; gene; ss.
XX
XX Homo sapiens.
OS
XX WC200270539-A2.
XX
XX 12-SEP-2002.
PD
XX 05-MAR-2002; 2002WO-US05095.
PF
XX 05-MAR-2001; 2001US-0799451.
PR
XX (HYSE-) HYSEQ INC.
PA
XX Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
PI Wehrman T, Wang J, Wang D, Drmanac R;
XX
XX WPI; 2002-759812/82.
DR
XX P-PSDB; ABP69127.
XX
XX New polynucleotides comprising sequences assembled from expressed
PT sequence tags (ESTs), useful for treating cell-proliferative,
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or
PT platelet or coagulation disorders -
XX
XX Claim 1; SEQ ID NO 226; 1012pp + Sequence Listing; English.
PS
XX The invention relates to an isolated polynucleotide (I) comprising a
CC nucleotide sequence selected from any of 948 sequences
CC (AB211119-AB212066) or their mature protein coding portion, active domain
CC coding protein or complementary sequences. The polynucleotides are useful
CC for identifying expressed genes or for physical mapping of human genome.
CC The encoded polypeptides (ABP68902-ABP69849) are useful as molecular
CC weight markers, as a food supplement, for generating antibodies, in
CC medical imaging, screening and diagnostic assays and for treating
CC cell-proliferative disorders (cancer), neurodegenerative diseases
CC (Parkinson's or Alzheimer's disease), autoimmune diseases (multiple
CC sclerosis, diabetes, lupus) genetic disorders, myeloid or lymphoid
CC disorders, platelet or coagulation disorders, wound, burns, incision,
CC ulcers, liver or lung fibrosis, infections (bacterial, viral, fungal,
CC parasitic), arthritis, etc.
CC Note: the sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1309 BP; 287 A; 384 C; 365 G; 273 T; 0 other;

Query Match 13.8%; Score 156.4; DB 24; Length 1309;
Best Local Similarity 53.6%; Pred. No. 2.2e-25;
Matches 437; Conservative 0; Mismatches 346; Indels 33; Gaps 4;

Qy 153 CCGCTGTGTGGCGACAGCGAGGCTTCTGACCGCAACTGACGCGCGCGCGACCC 212
Db 256 CGCACCCAGTGTGCGCGGACTTCCCGTGGAGTGGACCCCGACTGGTCCCCAGCCC 315

```
QY 213 CGAGCCCGCCGCGCCCTCGCAGCTATCCGAGTGTGGCGCTCTTCGGGGGCGCTGCTGCG 272
DB 316 GGCACAGGCTCGGGCGCGCGCGCTCGCGAGCTTCTTCGGGGGCGCTTCTGCG 375
QY 273 CGCGCGCAGCTGCTCAAGCGCTGCAAGCAGGCGCTGCGAGCTTCCGCGAGTCCCGAGCC 332
DB 376 TCGGCTGCTGCTGCGCGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 426
QY 333 CAGCGCGAGTGTGCGCGAGCTTCCAGCGCGCGCGAGCGCTTACAGTTCCTGCGAGTTCGC 392
DB 427 CAGCGAGAGATGGAGCTGGAGTTCGGCAAGCGGAGCGCGCTTACACTACCTGAGTTCGC 486
QY 393 TTACTTCAAGCAATAATCTCCCAAGCATCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 452
DB 487 CTACTTCAAGATCAACAGTTGGAGAAAGCTGTTGCTGCGAGCACACACCTTCTTCGTGGG 546
QY 453 GCATCTCATGACGAATGATGAAGAGAAATGATGATGATGATGATGATGATGATGATGATGATG 511
DB 547 CAATCTGAGCAGATGGAATGCGAGCACTAGACTATTAACCAACCATGCTGAGT 606
QY 512 --CGAGGACATGATTAAGACCTGGAACCAAGCTATATGAAGCGCTGCTGCTGCTGCTGCTG 569
DB 607 GAAGAGCGCGCTTCAAGGATCTGAGCTCAACCGCATATGCAAGAAATTCGACTGGG 666
QY 570 AGTGGGCGATACAGCGTGAGACTGGAGAACATGATGATGATGATGATGATGATGATGATGATG 629
DB 667 AGTGGGCTTACTCAGAGGAAACAGCCACAGGAAGCTGTGCGCCACCTAGAGCGCGGCT 726
QY 630 TCCGAGCTTCTCAAGCGCTTTCAGAGTGTCTGCGAGCTGCGAGGCTTCCAGGGAGAT 689
DB 727 GCAAGATACCTTGTGGCTATGAGGAGTGGCGCTCTGCGAAGGCGCTATGACTA 786
QY 690 CAAGACTTCAA-----GGATTTCTACCTTTCATGAGAGAACT---CACCCAGTTATAGG 731
DB 787 CGATGGCTTACAACTACCTTGTAGTCAAGCGTGAGCTTCTTCAGGCGCATCAGATCATTA 846
QY 732 TGTGAAGTCTGGAATCAAAATACAGTGTGAAGAGAACT---CACCCAGTTATAGG 788
DB 847 CATCAGTCTCTCACTGTAGCAGAACTGTGTACGAGGCTTCTTCCACCAAGTCG 906
QY 789 AGGTATCCGTTGAGAAATTTGGTGGCPACCATGATATCATGATGATGATGATGATGATGATGAT 848
DB 907 AGAGAAGCGCTTGAAGACTTCTCCCATCCATGATGATGATGATGATGATGATGATGATGATGAT 966
QY 849 TAAGTTGAACGCTGAGAGATGAGCGCGCGCTGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 908
DB 967 TAACATGGGAATATATACAGGCTGTTGAATGTCGAAGCACTATCTTCTTCTTCTTCTTCTT 1026
QY 909 GAATGACAAGTCTATGAGCAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 944
DB 1027 CAATGACGAGTGTATGAACCAAAATTTGGCTTATTA 1062

RESULT 10
ID ABZ11345
XX ABZ11345 standard; cDNA; 2152 BP.
AC ABZ11345;
XX ABZ11345;
DT 20-JAN-2003 (first entry)
DE Human polynucleotide SEQ ID NO 227.
KW Human; genome mapping; gene therapy; food supplement; virus; fungus;
KW cell-proliferative disorder; neurodegenerative disease; bacterial;
KW Parkinson's disease; Alzheimer's disease; autoimmune disease; multiple
KW sclerosis; diabetes; genetic disorder; wound; burn; infection;
KW arthritis; cyclostatic; immunomodulator; nontropic; neuroprotective;
KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
KW haemostatic; vulnary; fungicide; antibacterial; virucide; protozoacide;
KW antiarthritic; gene; ss.
```

```
OS Homo sapiens.
XX
PN WO200270539-A2.
XX
XX 12-SEP-2002.
PD
XX 05-MAR-2002; 2002WO-US05095.
PF
XX 05-MAR-2001; 2001US-0799451.
PR
XX (HYSE-) HYSEQ INC.
PA
XX Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
PI Wehrman T, Wang J, Wang D, Drmanac RT;
XX
DR WPI; 2002-759812/82.
DR P-PSDB; ABP69128.
XX
XX New polynucleotides comprising sequences assembled from expressed
PT sequence tags (ESTs), useful for treating cell-proliferative,
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or
PT platelet or coagulation disorders
XX
PS Claim 1; SEQ ID NO 227; 1012pp + Sequence Listing; English.
XX
XX The invention relates to an isolated polynucleotide (I) comprising a
CC nucleotide sequence selected from any of 948 sequences
CC (AB21119-AB212066) or their mature protein coding portion, active domain
CC coding protein or complementary sequences. The polynucleotides are useful
CC for identifying expressed genes or for physical mapping of human genome.
CC The encoded polypeptides (ABP68902-ABP69849) are useful as molecular
CC weight markers, as a food supplement, for generating antibodies, in
CC medical imaging, screening and diagnostic assays and for treating
CC cell-proliferative disorders (cancer), neurodegenerative diseases
CC (Parkinson's or Alzheimer's disease), autoimmune diseases (multiple
CC sclerosis, diabetes, lupus) genetic disorders, myeloid or lymphoid
CC disorders, platelet or coagulation disorders, wound, burns, incision,
CC ulcers, liver or lung fibrosis, infections (bacterial, viral, fungal,
CC parasitic), arthritis, etc.
CC
CC Note: the sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2152 BP; 494 A; 636 C; 602 G; 420 T; 0 other;
XX
Query Match 13.8%; Score 156.4; DB 24; Length 2152;
Best Local Similarity 53.6%; Pred. No. 2.5e-25;
Matches 437; Conservative 0; Mismatches 346; Indels 33; Gaps 4;
QY 153 CGGCTTGTGCGCGAGCAGCGAGGCGCTTCTGCGCACCGCAACTGCAGCGCGCGCGCGCGCC 212
DB 256 CCGCACCAGTGTGCGCGCGAGTTCCTGCGAGCTGGAGCTGGACCCCGAGTGGTCCCCAGGCC 315
QY 213 CGAGCCCGCGCGCGCTGCGCGAGCTATCCGAGTGTGCGCTTTCGGGGGCGCTGCTGCG 272
DB 316 GGCACAGGCTCGGGCGCGCGCGCTGCGAGCTTCTTCGGGGGCGCTTCTGCG 375
QY 273 CGCGCGCAGCTGCTCAAGCGCTGCAAGCAGGCGCTGCGAGCTTCCGCGAGTCCCGAGCC 332
DB 376 TCGGCTGCTGCTGCGCGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 426
QY 333 CAGCGCGAGTGTGCGCGAGCTTCCAGCGCGCGCGAGCGCTTACAGTTCCTGCGAGTTCGC 392
DB 427 CAGCGAGAGATGGAGCTGGAGTTCGGCAAGCGGAGCGCGCTTACACTACCTGAGTTCGC 486
QY 393 TTACTTCAAGCAATAATCTCCCAAGCATCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 452
DB 487 CTACTTCAAGATCAACAGTTGGAGAAAGCTGTTGCTGCGAGCACACACCTTCTTCGTGGG 546
QY 453 GCATCTCATGACGAATGATGAAGAGAAATGATGATGATGATGATGATGATGATGATGATGATG 511
DB 547 CAATCTGAGCAGATGGAATGCGAGCACTAGACTATTAACCAACCATGCTGAGT 606
```

QY 512 --CCGAGACTACATTAAAGACCTGGAAACCAAGTCAATCAAGCTGTTCATCCGAGC 569
 Db 607 GAAGAGGCGGACCTCAAGGATCTTGAGACTCAACCCATATGCAAGAAATTCGACTGGG 666
 QY 570 AGTCCGGGCATACAAACGCTGAGACTGAGAACTCCATCACAGACATGGAGTCCCT 629
 Db 667 AGTCCGACTTACTCAGAGHACAGCCACAGAAAGCTGTGCCCCACCTAGAGGGGGCT 726
 QY 630 TCCGAGCTTCTTCAAGCCCTTTTACGAGTGTCTCGCAGCTGCGAGGGTTCCAGGAGAT 699
 Db 727 GCAAGAAATCTTTGTGGCTATGAGGAGTGCCTGCGCCCTCTGCGAAGGGCCCTATGACTA 766
 QY 690 CAAGGACTTCAA-----GGATTCTACTTTCATAGCAGATCATTA 731
 Db 787 CGATGGCTACAACTACCTTGAGTAAACGCTGACCTTCCAGGCCATCACAGATCATTA 846
 QY 732 TGTAGAAGTCTGGAATGCAAAATACAGTGTGAAGAAACCT---CACCCCACTTATAGG 788
 Db 847 CATCCAGTCTCACTGTAGCAGAACTGTGTACAGAGCTTCTTCCACCCCAAGTCG 906
 QY 789 AGGTATCCGGTTGAGAAATTTGGGTACATGTATCATATCTTGAGTTTGCCTATTA 848
 Db 907 AGAGAAGCCCTTTGAAGACTTCTCCCATCGCATTAATATATCTGCAGTTTGCCTACTA 966
 QY 849 TAAGTTGAACGACCTGAGAAATGACGCCCTGTGCGAGTCAGTATCTGTCTTTGATCA 908
 Db 967 TAACATTTGGGAATATACAGAGCTGTGTGAATGCGCAAGACTATCTTCTTCTTCCC 1026
 QY 909 GAATGACAAGTCTATGAGCAGAACCTTGCTGATTA 944
 Db 1027 CAATGACGAGTGTATGAACCAAAATTTGGCCTATTA 1062

RESULT 11
 AAS17573
 ID AAS17573 standard; cDNA; 2211 BP.
 XX AC AAS17573;
 XX AC AAS17573;
 XX DT 26-FEB-2002 (first entry)
 XX DE DNA encoding novel secreted protein #2.
 XX KW Secreted protein; cytostatic; immunosuppressive; vulnery; vaccine;
 XX KW anti-inflammatory; neuroprotective; nephrotropic; cardiovascular;
 XX KW human; cancer; autoimmune disease; wound healing disorder; infection;
 XX KW haematopoietic disorder; inflammatory disorder; infertility;
 XX KW neurological disease; psychiatric disease; cardiovascular disease;
 XX KW respiratory disease; renal; gastrointestinal; ss.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 XX FT CDS 1..2211
 XX FT /*tag= a
 XX FT /product= "Human secreted protein"
 XX PN WO200179454-A1.
 XX XX 25-OCT-2001.
 XX XX 11-APR-2001; 2001WO-US11797.
 XX XX 13-APR-2000; 2000US-196603P.
 XX XX 24-APR-2000; 2000US-199417P.
 XX XX (SWIK) SMITHKLINE BEECHAM CORP.
 XX XX (SWIK) SMITHKLINE BEECHAM PLC.
 XX XX Agarwal P, Murdock PR, Rizvi SK, Smith RF, Xiang Z;
 XX XX WPI; 2002-061975/08.

P-PSDB; AAU09861.
 XX New secreted proteins or polypeptides, useful for treating e.g. cancer,
 PT autoimmune diseases, wound healing disorder, infections, haematopoietic
 PT disorders, inflammatory disorders, infertility, cancer
 XX Claim 2; Page 35-36; 92pp; English.
 XX The invention relates to an isolated novel secreted polypeptide (I) and
 CC polynucleotide (II). (I) and (II) are useful for treating cancer,
 CC autoimmune diseases, wound healing disorder, infections, haematopoietic
 CC disorders, inflammatory disorders, infertility, neurological and
 CC psychiatric diseases, cardiovascular diseases, respiratory diseases,
 CC renal diseases, or gastrointestinal diseases. These may also be used to
 CC treat diseases, abnormalities and disorders caused by abnormal
 CC expression, production, function and/or metabolism of the genes, as
 CC vaccines for inducing immunological response in a mammal, and in
 CC screening methods for detecting the effect of added compounds on the
 CC production of mRNA and polypeptide in cells. The polypeptides can be used
 CC as immunogens to produce antibodies immunospecific for the polypeptides,
 CC and to identify membrane-bound or soluble receptors. The polynucleotides
 CC may be used as diagnostic reagents, in chromosome localisation studies,
 CC and in tissue expression studies. The present sequence represents the
 CC coding sequence of novel human secreted protein #2.
 XX
 SQ Sequence 2211 BP; 512 A; 649 C; 612 G; 438 T; 0 other;
 Query Match 13.8%; Score 156.4; DB 24; Length 2211;
 Best Local Similarity 53.6%; Pred. No. 2.5e-25;
 Matches 437; Conservative 0; Mismatches 346; Indels 33; Gaps 4;
 QY 153 CGCCTTGTCTGCGGACAGCAGGAGGCTTCTGCGACCGAAGTGCAGCGCGCGCGAGCC 212
 Db 225 CCGACCCAGTGTGCGCGCGGAGCTTCCCGTGGAGCTGAGACCCGAGTGTCTCCGAGCC 284
 QY 213 CGAGCCCGCGCGGCTTCGCGAGCTATCCGAGCTGCGCTTTCGCGGCGCTGCTGCG 272
 Db 285 GCGCCAGGCTTGGCGCGCGCGCGCTGCGGAGCTGAGCTTCTTGGGGGCGCTTCTGCG 344
 QY 273 CGCGCGCACTGCTCAAGCGCTGCGAGCGGCTGCGAGCGCTTCCGCGAGTCCGAGCC 332
 Db 345 TCGCGCTGCGCTGCGCGCGCTGCGCTGCGCGCGCGCGCGCGCGCTC-----GCT 395
 QY 333 CAGCGCGAGTGTCTGCGGAGCTTCCAGCGCGCGGAGCGCTTACAGTTCCTGCGAGTTCGC 392
 Db 396 CAGCGAGAGATGAGCTGGAGTTCGCGAAGCGGAGCGCTTACACTACCTGCGAGTTCG 455
 QY 393 TTACTTCAAGCAATAATCTCCCAAGCATCGCGCTGCTCACACCTTTCTACTGAA 452
 Db 456 CTACTTCAAGATCAACAAGTTGGAGAAAGCTGTTGCTGCGAGCACACACCTTCTTCTG 515
 QY 453 GCATCTGATGACGAATGATGAAGAGAACTGATGATGATGATGATGATGATGATGATG 511
 Db 516 CAATCTGAGCAGATGGAATGCGAGAGACCTAGACTATTTACCAACCATGCTGAGT 575
 QY 512 --CCGAGACTACATTAAAGACCTTGAAGCAACCAAGTCAATGAAAGCGCTTTCATCCGAGC 569
 Db 576 GAAGGAGCGGACTTCAAGGATCTTGAAGCTCAACCCCATATGCAAGAAATTCGACTGGG 635
 QY 570 AGTCCGGGCATACAAACGCTGAGAACTGGAGAACTCCATCACAGACATGGAGTGGCGCT 629
 Db 636 AGTCCGACTTCTCAGAGGAAACAGCCACAGAAAGCTGTGCCCCACCTAGAGGGGGCT 695
 QY 630 TCCGAGCTTCTTCAAGCCCTTTTACGAGTGTCTCGCAGCTGCGAGGGTTCACAGGAGAT 689
 Db 696 CAAAGATATCTTTGTGGCTATGAGGAGTGCCTGCGCGCTCTGCGAAGGGCCCTATGACTA 755
 QY 690 CAAGGACTTCAA-----GGATTCTACTTTCATAGCAGATCATTA 731
 Db 756 CGATGGCTACAACTACCTTGAGTAAACGCTGAGCTTCTTCCAGGCCATCACAGATCATTA 815
 QY 732 TGTAGAAGTCTGGAATGCAAAATACAGTGTGAAGAAACCT---CACCCCACTTATAGG 788

816 CATCAGGCTCTCAACTGTAAAGAGAGAGTGTGTACAGGAGTGTCTCCACCAAGTCG 875
 789 AGGCTATCCGGTGTGAGAAATTTGGTACCAATGATATCACTTACCTGAGTTGGCTATTA 848
 876 AGAGAGCCCTTTGAAGACTTCTCCCATCGCATTAATTAATATCTGCGAGTTGGCTACTA 935
 849 TAAGTTGAACGCTGTGAAGATGAGAGCCCTGTGAGTCAAGTATCTGCTCTTTGATCA 908
 936 TAACATTTGGGAATTAACACAGGCTGTGAATGTGCCAAGACCTATCTCTCTCTCC 995
 909 GAATGACAGGCTCATCGACAGCAAGCTGTGTGATTA 944
 996 CATGACGAGGTGATGAACCAAAATTTGGCTATTA 1031

RESULT 12

ABZ11346
 ID ABZ11346 standard; cDNA; 2242 BP.

AC ABZ11346;

XX 20-JAN-2003 (first entry)

DE Human polynucleotide SEQ ID NO 228.

Human; genome mapping; gene therapy; food supplement; virus; fungus;
 cell-proliferative disorder; neurodegenerative disease; bacterial;
 Parkinson's disease; Alzheimer's disease; autoimmune disease;
 multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
 arthritis; cystic; immunomodulator; nontropic; neuroprotective;
 antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
 haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
 antiarthritic; gene; ss.

XX Homo sapiens.

XX WO200270539-A2.

XX 12-SEP-2002.

XX 05-MAR-2002; 2002WO-US05095.

XX 05-MAR-2001; 2001US-0799451.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
 XX Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
 XX Wehrman T, Wang J, Wang D, Drmanac RT;

XX WPI; 2002-759812/82.

XX P-PSDB; ABP69129.

XX New polynucleotides comprising sequences assembled from expressed
 XX sequence tags (ESTs), useful for treating cell-proliferative,
 XX neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or
 XX platelet or coagulation disorders

XX Claim 1; SEQ ID NO 228; 1012pp + Sequence Listing; English.

XX The invention relates to an isolated polynucleotide (I) comprising a
 XX nucleotide sequence selected from any of 948 sequences
 XX (ABZ1119-ABZ12086) or their mature protein coding portion, active domain
 XX coding protein or complementary sequences. The polynucleotides are useful
 XX for identifying expressed genes or for physical mapping of human genome.
 XX The encoded polypeptides (ABP6902-ABP69849) are useful as molecular
 XX weight markers, as a food supplement, for generating antibodies, in
 XX medical imaging, screening and diagnostic assays and for treating
 XX cell-proliferative disorders (cancer), neurodegenerative diseases
 XX (Parkinson's or Alzheimer's disease), autoimmune diseases (multiple
 XX sclerosis, diabetes, lupus) genetic disorders, myeloid or lymphoid
 XX disorders, platelet or coagulation disorders, wound, burns, incision,
 XX ulcers, liver or lung fibrosis, infections (bacterial, viral, fungal,

CC parasitic), arthritis, etc.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 2242 BP; 518 A; 655 C; 625 G; 444 T; 0 other;

Query Match 13.8%; Score 156.4; DB 24; Length 2242;
 Best Local Similarity 53.6%; Pred. No. 2.5e-25;
 Matches 437; Conservative 0; Mismatches 346; Indels 33; Gaps 4;

QY	153	CGCTTGTCTGCGGACAGGAGGCTTCTGCGCACCGCACTGACGCGCGCGCAGCC 212
DB	256	CGCACCCAGTGTGCGCGGACTTCCGTGGAGTGGACCCGACTGCTGCTCCCGAGCC 315
QY	213	CGAGCCCGCGCGGCTCTGCCAGTATCCCGAGTGTGCGCTTTCGGGGGCTCTGCG 272
DB	316	GGCCAGGCTCGGGCGCGCGCTGCGGACCTGAGCTTCTTCGGGGGCTTCTGCG 375
QY	273	CGCGGCGACTGCTCAAGCGTGTGCAAGCAGGCGCTGCGAGCTTCGCGCATGCCAGCC 332
DB	376	TGCGCTGCTGCTGCGCGCTGCTGCGCGCGCGCTGCGCGCGCGCGCGCTC-----GCT 426
QY	333	CAGCCGCGAGTGTCTGCGGAGCTTTCAGCGCGCGCGCGCGCTACAAAGTTCTCTGAGTTGCG 392
DB	427	CAGCGAAGAGTGTGAGCTGAGTTCGCGAAGCGGAGCGCTTACAACTACTCTCAGTGTGC 486
QY	393	TACTTCAAGGCAATATCTCCCAAGCCATCGCGCTGTCTACACCTTCTACTGAA 452
DB	487	CTACTTCAAGATCAACAAGTTGAGAAAGCTGTGCTGCGACACACCTTCTCTGTTGG 546
QY	453	GCATCTGTATGAGAAATGATGAAGAGGAACATGGCATATTAAGAGCTGCTCTGGTG- 511
DB	547	CAATCTGTAGCATGTGAAATGCAGCAGAACCTAGACTATTACCAACCATCTCTGGAGT 606
QY	512	--CCGAGGACTACATTAAGACCTGGAACCAAGTATATGAAGCCTTTCATCCGAGC 569
DB	607	GAAGGAGGCGGACTTCAAGGATCTTGAGACTCAACCCATATGCAAGAATTTGCGACTGG 666
QY	570	AGTGGCGGCATACACGGTGAGAACTGGAGAACATCCATCAGACATGAGCTGCGCCT 629
DB	667	AGTGGCACTTCTAGAGGAACAGCCACAGAAAGTGTGCGCCCTTAGAGGCGCGCT 726
QY	630	TCCGAGCTTCTTCAAGCCCTTTAGAGTGTCTCGCAGCTGCGAGGTTCCAGGGAGAT 689
DB	727	GCAAGAATACTTTGTGGCTATGAGAGTGTGCGCTCTCGAAGGCGCTTATGACTA 786
QY	690	CAGGACTTCAA-----GGATTCTACCTTTGCATAGAGATCATTA 731
DB	787	CGATGCTCAACTACCTTTGAGTACAAACGCTGACCTTTCAGGCGCATCACAGATCATTA 846
QY	732	TGTAGAAGTTCTGGAATGCAAAATACAGTGTGAAGAGAACCT---CACCCAGTTATAGG 788
DB	847	CATCCAGTCTCACTGTAAAGACAGTGTGTACGAGGCTTGTCTCCACCCAGTGC 906
QY	789	AGCTATCCGGTTGAGAAATTTGTGGCTACCATGATCATTTACTTTCAGTTTGCCTATTA 848
DB	907	AGAGAAGCCCTTTGAGACTTCTCCCATCGCATTAATTAATCTGCAAGTTTGCCTACTA 966
QY	849	TAAGTTGAACGACCTGAAGAATGCGCCCTGTGCGAGCTATCTGCTCTTTGATCA 908
DB	967	TAACATTTGGAATTAACAGAGCTGGTGAATGTGCCAAGACCTATCTCTCTTCTCC 1026
QY	909	GAATGACAGGCTCATGACAGCAAACTGGTGATTA 944
DB	1027	CAATGACGAGGTGATGAACCAAAATTTGGCTTATTA 1062

RESULT 13

AAH14513
 ID AAH14513 standard; cDNA; 2524 BP.
 XX
 AC AAH14513;

XX 26-JUN-2001 (first entry)
 XX Human cDNA sequence SEQ ID NO:12044.
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 XX Homo sapiens.
 OS EP1074617-A2.
 PN 07-FEB-2001.
 XX 28-JUL-2000; 2000EP-0116126.
 XX 29-JUL-1999; 99JP-0248036.
 XX 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX (HELI-) HELIX RES INST.
 XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 DR WPI; 2001-318749/34.
 XX Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX Claim 8; SEQ ID 12044; 2537pp + CD ROM; English.
 XX The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
 CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX Sequence 2524 BP; 583 A; 747 C; 704 G; 490 T; 0 other;
 SQ Query Match 13.8%; Score 156.4; DB 22; Length 2524;
 Best Local Similarity 53.6%; Pred. No. 2.6e-25;
 Matches 437; Conservative 0; Mismatches 346; Indels 33; Gaps 4;
 QY 153 CGCTGTGTCGCGCAGCAGGAGCTTCTGCGCCACCGCACTGACGCGCGCGCAGCC 212
 DB 240 CGCACCACTGTCGCGCCGACCTTCCGTCGGAGTGACCCCGACTGTCCTCCCGAGCCC 299
 QY 213 CGAGCCCGCGCGGCTCCGACGATATCCGAGCTGCGCTCTTCGGGGGCGCTGCTGCG 272
 DB 300 GCGCCAGGCTCGGGGCGCGCGCCCTGCGGACCTGAGCTTCTTCGGGGGCGCTTCTCGG 359
 QY 273 CCGCGGCGACTGCCTCAAGCGCTGCGAGCAGGCGCTGCCAGCTTCGCGCAGTCCAGCC 332

DB 360 TCGCGCTGCTGCTGCGCGCTGCTGCGCGCGCGCGCGCGCGCGCTC-----GCT 410
 QY 333 CAGCGCGAGGTGCTGCGGACTTCCAGCGCGCGGAGCCCTACAAGTTCCTGCTGAGTTCGC 392
 DB 411 CAGGAAGAGTGGAGCTGGAGTTCGGAAGCGAGCCCTACAACCTACCTGCGAGGTGCG 470
 QY 393 TTACTTCAAGGCAATAATCTCCCAAGCCATCGCGCTGCTCAGACCTTTCTACTGAA 452
 DB 471 CTACTTCAAGATCAACAAGTTGGAGAAAGCTGTTGTCGAGCAGACACACCTTCTTCGTGG 530
 QY 453 GCATCTGATGACGAATGATGAAGGAGACATGCGCATATATATAAGAGCCTGCTGCTG- 511
 DB 531 CAATCTCTGAGCAGATGGAATGCGAGAGAACCTAGACTATTACCAAAACCATGCTCGAGT 590
 QY 512 --CGAGGACTACATTAAAGACCTGGAACCAAGTCAATATGAAAGCTGTTTCATCCGAGC 569
 DB 591 GAAGGAGCGGACTTCAAGGATCTTGAGACTCAACCCCATATGCAAGAAATTCGACTGGG 650
 QY 570 AGTGGGGCATACAACGGTGAAGAACTGGAGAACATCCATCAGACAGATGGAGCTGGCCCT 629
 DB 651 AGTGGCACTCTACTCAGAGGAGACAGCCACAGGAAGCTGTGCCCCACCTAGAGCGCGCT 710
 QY 630 TCCGACTTCTTCAAGCCCTTTTACGAGTGTCTCGCAGCTGCGAGGTTCCAGGAGAT 689
 DB 711 GCAAGAAATCTTTGTGCTTATGAGGAGTCCCTGCGAGGCGCTTATGACTA 770
 QY 690 CAAGGACTTCAA-----GGATTCTTACCTTTCCATAGCAGATCATTA 731
 DB 771 CGATGGCTCAACTACCTTGAGTACAACGCTGACCTTCTCCAGGCCATCAGACATCATTA 830
 QY 732 TGTAGAAGTCTGGAATGCAAAATACAGTGTGAAGAACCT---CACCCAGTTATAG 788
 DB 831 CATCCAGGCTCTCACTGTAAAGCAGAACTGTGTCAGGAGCTTGTCCACCCCAAGTCG 890
 QY 789 AGGCTATCCGGTTGAGAAATTTGTGCTACCATGTATCATTTACTTGCAGTTTGCCTATTA 848
 DB 891 AGAAGAGCCCTTTGAAGACTTCTCCCATCGCATTAATATATCTGAGTTTGCCTACTA 950
 QY 849 TAAGTTGAAGCACTGAAGAAATGCAAGTCCAGCCCTGTCGAGTCAGTATCTGCTTTGATCA 908
 DB 951 TAACTTGGGAATTATACACAGGCTGTTGAATGTGCCAAGACCTATCTTCTCTCTTCCC 1010
 QY 909 GAATGACAAGCTCATGCGCAGCAGAACTGGTGTATTA 944
 DB 1011 CAATGACAGGCTGATGAACCAAAATTTGCGCTATTA 1046
 RESULT 14
 AAF93800
 ID AAF93800 standard; cDNA; 2563 BP.
 AC AAF93800;
 XX 23-MAY-2001 (first entry)
 DT Human cDNA encoding a membrane or secretory protein clone PSEC0109.
 XX Human; secretory protein; membrane protein; vaccine; gene therapy;
 KW rheumatoid arthritis; diabetes; ss.
 XX Homo sapiens.
 OS EPI067182-A2.
 PN 10-JAN-2001.
 XX 07-JUL-2000; 2000EP-0114090.
 PF 08-JUL-1999; 99JP-0194179.
 PR 11-JAN-2000; 2000JP-0118775.
 PR 02-MAY-2000; 2000JP-0183766.
 XX

PA (HELI-) HELIX RES INST.
 XX Ota T, Isogai T, Nishikawa T, Kawai Y, Sugiyama T, Hayashi K;
 PI WPI: 2001-092989/11.
 XX P-PSDB; AAB88373.
 DR
 XX Nucleic acids encoding secretory proteins/membrane proteins, useful in
 PT gene therapy or as candidate target molecules in drug development -
 XX Claim 1; SEQ ID 113; 609pp + CD ROM; English.
 PS
 XX This invention relates to nucleic acid sequences AAF93744 - AAF93916
 CC which encode human secretory or membrane proteins represented by
 CC AAB88317 - AAB88419. Included in the invention are primers
 CC AAF93917 - AAF94295 and AAF62232 - AAF62235 which are used to isolate the
 CC cDNA sequences of the invention. The invention also includes methods for
 CC the production of antibodies directed against the proteins, and cDNA
 CC sequences, which can be used in vaccines. The polynucleotide sequences
 CC can be used in gene therapy. The polynucleotide sequences and the
 CC proteins they encode may be used in the prevention, treatment and
 CC diagnosis of diseases associated with inappropriate secretory
 CC protein/membrane protein expression. The nucleic acids and complementary
 CC sequences may also be used as DNA probes in diagnostic assays
 CC (e.g. polymerase chain reactions (PCR)) to detect and quantitate the
 CC presence of similar nucleic acid sequences in samples. They may also be
 CC used to study the expression and function of secretory proteins/membrane
 CC polypeptides and their role in metabolism. The polypeptides may be used
 CC as antigens in the production of antibodies against them and in assays to
 CC identify modulators (agonists and antagonists) of expression and
 CC activity. The antibodies and antagonists may also be used as therapeutic
 CC agents to down regulate expression and activity. The antibodies may also
 CC be used as diagnostic agents for detecting the presence of the
 CC polypeptides in samples (e.g. by enzyme linked immunosorbent assay
 CC (ELISA). Examples of diseases which may be treated include rheumatoid
 CC arthritis and diabetes.
 XX
 SQ Sequence 2563 BP; 589 A; 753 C; 713 G; 508 T; 0 other;
 Query Match 13.8%; Score 156.4; DB 22; Length 2563;
 Best Local Similarity 53.6%; Pred. No. 2.6e-25;
 Matches 437; Conservative 0; Mismatches 346; Indels 33; Gaps 4;
 QY 153 CCGCTTGTCTGCGACAGCAGGCGCTTCTCCACCGCAACTGCAGGCGCGCGCGCAGGC 212
 DB 255 CCGACCCAGGTGCGCCCGACTTCCGCTGGAGCTGGACCCCGACTGCTCCCGACGCC 314
 QY 213 CGAGCCGCGCGCGCTCGCAGCTATCCGAGCTGCGCTCTTCGGGGCGCTGCTGCG 272
 DB 315 GCGCCAGGCGCTCGGGCGCGCGCGCTGCGCGACCTGAGCTTCTTCGGGGCGCTTCTGCG 374
 QY 273 CCGCGCGCACTGCTCAAGCGCTGCAAGCGGCGCTGCGAGCGCTTCCGCCAGTCCCGCAGCG 332
 DB 375 TCGCGCTGCTGCTGCGCGCTGCTGCGCGCGCGCGCGCGCCATC-----GCT 425
 QY 333 CAGCGCGAGGTGCTGCGGACTTCCAGCGCGCGCGCGCTCAAGTCTCTGAGTTCGC 392
 DB 426 CAGCGAAGAGATGAGTGGAGTTCGCGAAGCGCGCGCGCGCTTACCACTACCTGCGAGTGC 485
 QY 393 TTACTTCAGGCAATATCTCCCAAGCGCATGCGCGCTGCACACCTTCTACTGAA 452
 DB 486 CTACTTCAAGATCAACAGTTGGAGAAAGCTGTGCTGCGAGCACACCTTCTCTGTTGG 545
 QY 453 GCATCTTGATGACAAATGATGAAGAGAGAAATGGCATATTATAAGAGCGCTGCGTGTG- 511
 DB 546 CAATCTTGAGCACATGGAATGACAGAGAACTAGACTATTACCAACCATGCTGGAGT 605
 QY 512 --CCGAGGACTACATTAAAGCTTGGAAACCAAGTCTATGAAGCTGTTCATCCGAGC 569
 DB 606 GAAGAGAGCGGACTTCAAGGATCTTGAGACTTCAACCCCATATGCAAGAAATTGAGCTGGG 665
 QY 570 AGTGGGCGCATACAAACGCTGAGAACTGGAGAAATCCATCCATCAGACATGAGCTGGCGCT 629

DB 666 AGTGGGACTCTACTCAGAGGAAACAGCCACAGAAAGCTGTGCCCACTAGAGGGCGGCT 725
 QY 630 TCCGACTCTTCAAGGCTTTTACGAGTGTCTCGAGCTGCGAGGCTTCCAGGGAGAT 689
 DB 726 GCAAGATATCTTGTGGCTATGAGAGTGCCTGCGAGGGGCGCTTATGACTA 785
 QY 690 CAAGGACTTCAA-----GGATTCTTACCTTTCCATAGCAGATCATTA 731
 DB 786 CGATGGCTACAACTACCTTGTAGTACAAAGCTGACCTCTTCCAGGCGCATCAGATCATTA 845
 QY 732 TGTAGAGTCTTGGAAATGCAAAATACAGTGTGAAGAGAACT--CACCCAGTTATAGG 788
 DB 846 CATCCAGTCTCTCACTGTAAGCAGAACTGTGTCCAGGAGCTTCTTCCACCAAGTCG 905
 QY 789 AGGTATCCGGTTGAGAAATTTGGTGTACCATATCATTTACTTTCAGTTTGGCTATTA 848
 DB 906 AGAGAAGCCCTTGAAGACTTCTCCCATCGATTATAATATCTGCGAGTTTGGCTACTA 965
 QY 849 TAAGTTGAACCACTGAGATGAGATGAGCCCTGTCAGTGCAGTATCTGCTTGTGATCA 908
 DB 966 TAACATTGGGAATATATACAGGCTGTGAATGTGCCAAGACCTATCTTCTTCTTCCC 1025
 QY 909 GAATGACAAGGTCTATGCGAGCAGAACTGGTGTATTA 944
 DB 1026 CAATGACGAGGTGATGACCAAAATTTGGCTATTA 1061
 RESULT 15
 AAH14589
 ID AAH14589 standard; cDNA; 2585 BP.
 XX
 AC AAH14589;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human cDNA sequence SEQ ID NO:12193.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI: 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX
 PS Claim 8; SEQ ID 12193; 2537pp + CD ROM; English.
 CC
 CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination

of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AB92446 to AB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

Sequence 2585 BP; 590 A; 759 C; 723 G; 513 T; 0 other;

Query Match 13.8%; Score 156.4; DB 22; Length 2585;
Best Local Similarity 53.6%; Pred. No. 2.6e-25;
Matches 437; Conservative 0; Mismatches 346; Indels 33; Gaps 4;

Qy	153	CCGCTTCTCGCGCACAGCGAGGCTTCTGCCACCGCAACTGCGAGCGCGCGCGCGAGCC	212
Db	277	CCGCACCCAGTGTGCCCGCGACTTCCCGTGGAGCTGGACCCCGACTGTGTCCTCCCGAGCC	336
Qy	213	CGAGCCCGCGCGCGCTCCCGAGCTATCCCGAGCTGCGCTCTTCGGGGGCGCTGTGGG	272
Db	337	GGCCCGAGGCTCGGGCGCGCGCTCGCGCTGCGAGCTTCTTCGGGGGCGCTTCTGCG	396
Qy	273	CCGGCGCGCACTGCGCTCAAGCGCTGCAAGCAGGGGCTGCCAGCGCTTCGCCAGTCCCAGCC	332
Db	397	TGGGCTGCGCTGCGCTGCGCGCTGCTGCGCGCGCGCGCGCGCTC-----GCT	447
Qy	333	CAGCCCGAGGCTGCTGCGGACTTCCAGCGCGCGAGCCCTACAAGTTCCTGCACTGCG	392
Db	448	CAGCGAAGAGATGAGCTGGAGTTCGCGAAGCGAGCGCCCTACAACTACTGCGAGTGG	507
Qy	393	TTACTTCAAGGCAATATCTCCCAAGCCATCGCGGCTGCTCAGACCTTCTACTGAA	452
Db	508	CTACTTCAGATCAACAAGTTGGAGAAAGCTGTTGCTGAGCAGCACACCTTCTTGTGGG	567
Qy	453	GCATCTTGATGAGCAATGATGAAGAGGAACATGGCATATTAAGAGCGCTGCTGGTG-	511
Db	568	CAATCTTGAGCACATGGAATGAGCAGAGAACCTAGACTATTACCAACCATGCTCGAGT	627
Qy	512	--CGAGGACTACATTAAGACCTGGAACCAAGTCAATATGAAGCGCTGTTATCCGAGC	569
Db	628	GAAGGAGGCGGACTTCAAGGATCTTGAGACTCAACCCCATATGCAAGAATTTGACTGGG	687
Qy	570	AGTGGGGGCATACAAACGGTGAGAACTGGAGAACATCATCACAGACATGGAGTGGCCCT	629
Db	688	AGTGGGACTCTACTCAGAGGAAACAGCCACAGGAAGCTGTGCCCACTAGAGCGCGCT	747
Qy	630	TCCGAGCTTCTTCAAGCGCTTTACAGTGTCTCGAGCGCTGCGAGGGTTCAGGGAGAT	689
Db	748	GCAAGAATACTTTGTGGCTTATGAGGAGTCCCGTGGCCCTCTGCGAAGGGCCCTATGACTA	807
Qy	690	CAAGGACTTCAA-----GGATTCTACCTTTCCATAGCAGATCATT	731
Db	808	CGATGGCTAACACTCCTTGAGTACACGCTGACCTCTTCCAGGCCATCACAGATCATT	867
Qy	732	TGTAGAAGTTCTGGAATGCAAAATACAGTGTGAAGAGAACCT---CACCCCGAGTTATAGG	788
Db	868	CATCCAGGTCTCAACTGTAAAGCAGAACGTGTACGGAGCTTGTCCCAACCCCAAGTCG	927
Qy	789	AGGCTATCCGGTTGAGAAATTTGGGCTACCATGATCATCTACTTCAGTTTGCCTATTA	848
Db	928	AGAGAGCCCTTTGAGAGCTTCTCCCATCGCATTAATATCTGAGTTTGCCTACTA	987
Qy	849	TAAAGTTGAACGAGCTGAAGAAATGACGCCCGCTGTCAGTCTGCTCTTTGATCA	908

Db	988	TAACTTGGGAATTATACACAGGCTTTGAATGTGCCAAGACCTATCTTCTTCTTCCC	1047
Qy	909	GAATGACAAGGTCTATGTCAGCAGAACCTGGTGTATTA	944
Db	1048	CAATGACGAGGTGATGAACCAAAATTTGGCCTATTA	1083

Search completed: January 30, 2004, 08:32:19
Job time : 297.825 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2004, 08:13:15 ; Search time 406.168 Seconds
(without alignments)
10794.564 Million cell updates/sec

Title: US-09-729-674-1_COPY_63_1265

Perfect score: 1203

Sequence: 1 atggagcggggcgccggggg.....tgggaactggaggagaccagc 1203

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2434939 seqs, 1822278265 residues

Total number of hits satisfying chosen parameters: 4869878

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1202.2	99.9	3871	9	US-09-729-674-1
2	919	76.4	1810	12	US-10-108-260A-2058
3	467	38.8	2443	13	US-10-094-749-724
4	465.8	38.7	2347	10	US-09-962-832-113
5	465.8	38.7	2347	12	US-10-439-388-64
6	195.6	16.3	398	12	US-10-242-535A-8221
7	154.8	12.9	2600	14	US-10-045-815-3
8	153.8	12.9	2829	14	US-10-045-815-1
9	153.8	12.8	892	13	US-10-027-632-161474
10	153.8	12.8	892	13	US-10-027-632-161475
11	153.8	12.8	892	13	US-10-027-632-161476
12	153.8	12.8	892	14	US-10-027-632-161474
13	153.8	12.8	892	14	US-10-027-632-161475
14	153.8	12.8	892	14	US-10-027-632-161476
15	147.4	12.3	2322	14	US-10-045-815-7

16	147.4	12.3	2416	14	US-10-045-815-5
17	145.4	12.1	564	13	US-10-027-632-114858
18	145.4	12.1	564	14	US-10-027-632-114858
19	131.6	10.9	3396	13	US-10-094-749-921
20	126	10.5	618	10	US-09-879-536-215
21	109.2	9.1	2665	14	US-10-071-785-86
22	107.2	8.9	926	12	US-10-242-535A-9365
23	103.4	8.6	150	12	US-10-242-535A-1775
24	97.8	8.1	2753	10	US-09-728-952-28
25	97.8	8.1	2753	10	US-09-728-952-51
26	77	7.2	349	9	US-09-925-302-395
27	73	6.1	247	12	US-10-242-535A-4919
28	69.6	5.8	181	13	US-10-028-386-16058
29	69.6	5.8	549	13	US-10-028-386-16058
30	68.2	5.7	3897	13	US-10-023-386-2358
31	68.2	5.7	3897	14	US-10-021-657-7
32	60	5.0	1906	13	US-10-412-000-1
33	60	5.0	1906	14	US-10-021-657-1
34	59.8	5.0	2298	10	US-09-476-242-22
35	59.8	5.0	2298	10	US-09-476-242-23
36	59.8	5.0	2298	10	US-09-476-242-24
37	59.8	5.0	2310	10	US-09-476-242-3
38	59.8	5.0	2310	10	US-09-476-242-21
39	59.8	5.0	2316	10	US-09-476-242-4
40	59.8	5.0	2316	10	US-09-476-242-8
41	59.8	5.0	2322	10	US-09-476-242-5
42	59.8	5.0	2322	10	US-09-476-242-18
43	59.8	5.0	2322	10	US-09-476-242-19
44	59.8	5.0	2322	10	US-09-476-242-20
45	59.8	5.0	2328	10	US-09-476-242-6

ALIGNMENTS

RESULT 1

US-09-729-674-1
; Sequence 1, Application US/09729674
; Patent No. US20010039335A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Agostino, Michael J.
; APPLICANT: Steininger II, Robert J.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fectel, Kim
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 6055-64X
; CURRENT APPLICATION NUMBER: US/09/729,674
; CURRENT FILING DATE: 2000-12-04
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 283
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3871
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-729-674-1

Query Match 99.9%; Score 1202.2; DB 9; Length 3871;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCGGGCGCGGGGGCGGGCGGTAGCGCTGCTGTGGTGGCTGGCG


```

QY 601 CTGGAACCAAGTCATATGAAGCCCTGTTTCATCCGAGCAGTCCGGGCATCAACCGTGGAG 660
Db 624 CTGGAACCAAGTCATATGAAGCCCTGTTTCATCCGAGCAGTCCGGGCATCAACCGTGGAG 683
QY 661 AACTGGAGAACATCCATCAGACATGAGCTGGCCCTTCCCGATCTTCTTCAAGCCCTTT 720
Db 684 AACTGGAGAACATCCATCAGACATGAGCTGGCCCTTCCCGATCTTCTTCAAGCCCTTT 743
QY 721 TAGCAGTGTCTCGCAGCCCTGCCAGGCTTCCAGGGAGATCAAGGATCTTCAAGGATTTCTAC 780
Db 744 TAGCAGTGTCTCGCAGCCCTGCCAGGCTTCCAGGGAGATCAAGGATTTCTAC 803
QY 781 CTTTCCATAGCAGATCATATGATAGAGTTCTGGAATGCAAAATACAGTGTGAAGAGAAC 840
Db 804 CTTTCCATAGCAG----- 816
QY 841 CTCACCCAGTTATAGGAGGCTATCCGGTTGAGAAATTTGTGGCTACCATGTATCATTTAC 900
Db 817 ----- 816
QY 901 TTGCAAGTTTGCCTATTATAAGTTGAACGACCTGAAGAATGACGCCCTTGTGCAGTCAGC 960
Db 817 -----TGAACGACCTGAAGAATGACGCCCTTGTGCAGTCAGC 854
QY 961 TATCTGCTCTTTGATCAGATGACAGAGTGCATGCGAGCAACCTGGTGTATTACAGTAC 1020
Db 855 TATCTGCTCTTTGATCAGATGACAGAGTGCATGCGAGCAACCTGGTGTATTACAGTAC 914
QY 1021 CACAGGGACACTTGGGGCCCTCTCGGATGAGCACTTCCAGCCAGACCTGGAAGCAGTTTCAG 1080
Db 915 CACAGGGACACTTGGGGCCCTCTCAGATGAGCACTTCCAGCCAGACCTGGAAGCAGTTTCAG 974
QY 1081 TTCTTTAATGTACACACTCAGAAAGAGCTGTATGACTTTGCTTAAGGAAAAATATAATG 1140
Db 975 TTCTTTAATGTACACACTCAGAAAGAGCTGTATGACTTTGCTTAAGGAAAAATATAATG 1034
QY 1141 GATGATGATGAGGAGAGTGTGGAATATGGAATGATGCTCTTGGAACTGGAGGAGACC 1200
Db 1035 GATGATGATGAGGAGAGTGTGGAATATGGAATGATGCTCTTGGAACTGGAGGAGACC 1094
QY 1201 AGC 1203
Db 1095 AGC 1097

```

RESULT 3

```

US-10-094-749-724
; Sequence 724, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOKUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435

```

```

; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: Patentin ver. 2.1
; SEQ ID NO 724
; LENGTH: 2443
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-094-749-724

Query Match      38.8%; Score 467; DB 13; Length 2443;
Best Local Similarity 64.1%; Pred. No. 3,7e-126;
Matches 755; Conservative 2; Mismatches 402; Indels 18; Gaps 3;

QY 8 CCGGGCCCGGGGGCGCGGCTGCTAGCGTGTGTGCTGGCTGCGCGCTCGCG 67
Db 451 CCGGGCCCGGGGGCGCGGCTGCTAGCGTGTGTGCTGGCTGCGCGCTCGCG 510
QY 68 CCGGGCCCGGGGGCGCGGCTGCTAGCGTGTGTGCTGGCTGCGCGCTCGCG 127
Db 511 CCGGGCCCGGGGGCGCGGCTGCTAGCGTGTGTGCTGGCTGCGCGCTCGCG 570
QY 128 CGCTCGAGTCCGCTTACCGGCGCGCTGCGAGTACAGCGCGCGAGCACTG 187
Db 571 CGCTCGAGTCCGCTTACCGGCGCGCTGCGAGTACAGCGCGCGAGCACTG 630
QY 188 GCSTTGGCTTACCTTGGAGATCAGCGCTGCGGCTGCGAGCTTGTCTGCG 247
Db 631 CGCGCGCTTACCTTGGAGGCGCGCTGCGGCTGCGAGCTTGTCTGCGAG 690
QY 248 TCTGCCACCGCAACTGCGAGCG-----CGCGCGCGCGAGCGCGCGCG 298
Db 691 TCTGCCACCGCAACTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 750
QY 299 TCGCAGCTATC-----CGAGCTGCGCTCTTGGGGGCGCTGCTGCGCG 352
Db 751 CGCGAGCAGTGGGCTGCGAGCTGCGGCTCTTGGCGCGCTTCTGAGCGAG 810
QY 353 GCCTAAGCGCTCAAGCAGCGCGCTGCCAGCTTCCGCCAGTCCAGCGCGAG 412
Db 811 GCCTGCGCGCTCAAGCAGCGCGCTGCCAGCTTCCAGCTGCGCTTCCAGG 870
QY 413 TGCTGGCGACTTCCAGCGCGCGCGCTTCCAGCTTCTGAGTTCGCTTACTTCA 472
Db 871 TGCTGGCGACTTCCAGCGCGCGCTTCCAGCTTCTGAGTTCGCTTACTTCA 930
QY 473 CAAATAATCTCCCAAGCGCATCGCGCTGCTCACACCTTCTTACTGAAAGCA 532
Db 931 CTAACCGGCTGGAGAGCGGCTGGCGCGCGCTTACACCTTCTTCCAGAGGA 990
QY 533 ACGAAATGATGAAGAGGAACATGGCATATTATAA---GAGCCTGCTGTGCG 589
Db 991 ACGAGCTGACCGCAAGTATCTCACTACTATCAGGGGATGCTGGACGCTG 1050
QY 590 ACATTAAGACCTGGAAACCAAGTCATATAAGAACCTGTTTCATCCGAGCAG 649
Db 1051 CCCTCAGGACCTAGAGGCCAGCCCTACAGCGCTGTTCTCCGGGTGTGAAG 1110
QY 650 ACAACGGTGAAGACTGGAGAACATCCATCAGACATGAGCTGGCCCTTCCG 709
Db 1111 ACAACAGCGGGGATTTCCGAGCAGCACGAGAGCAGATGAGCGGGCTTGT 1170
QY 710 TCAAAGCCTTTTACGAGTGTCTCGAGCCTGGAGGTTCCAGGGAGATCAAGG 769
Db 1171 TGGCAGTCTTTTGGCCCGGCTGGCGCGCTTGAAGGGGCCCATGAGCAGG 1230
QY 770 AGGATTTCTACCTTTCCATAGCAGATCATATTATGTAAGTCTTGGAAATCA 829
Db 1231 AGGATTTCTACCGCGCCATAGCAGATCTCTTTGAGAGTCCCTTGGCAG 1290
QY 830 GTGAAGAGAACCTCACCCCGAGTTATAGAGGCTATCCGGTTGAGAAATTT 889

```

Db 1291 GTGAGGCAATTTGACCCCAATGTGGTGGCTACTTTCTGTGACAAAGTTCTGTGGCCACCA 1350
Qy 890 TGTATCATTTACTTCAGTTGGCTTATTAAGTTGACGACCTCAAGAAATGACAGCCCT 949
Db 1351 TGTACCACTACTTCAGTTGGCTTATTAAGTTGACGACCTCAAGAAATGACAGCCCT 1410
Qy 950 GTGCAGTCAGTATCTCTCTTTGATCAGAAATGACAAAGTTCATGACGACAACTGTGT 1009
Db 1411 GCGCGGCAGCTACATGCTCTTACCCCAAGACAGCGTCATGACGACAACTGTGT 1470
Qy 1010 ATTACCACTACCAAGGACACATGGGGCTCTCGATGACGACCTTCACCCAGACCTG 1069
Db 1471 ATTACCGGTTCCACCGGCTCTGCGGCTTGAAGAGGAGGACTTCCAGCCCGGAGG 1530
Qy 1070 AAGCAGTTCAGTTCTTTAATGTGACCACTCTCAGAAAGGAGCTGTATGATTTGCTAAG 1129
Db 1531 AGGCATGCTCTACCAACACAGACCCGAGCTGCGGAGCTGCTGGAGTTCAACCACA 1590
Qy 1130 AATATATATGATGATGAGGAGAAAGTTGTGA 1166
Db 1591 TGTACCTGCGATCGATGATGAGATGAGCTGGAGGA 1627

RESULT 4

US-09-962-832-113
; Sequence 113, Application US/09962832
; Patent No. US2002110821A1
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-74
; CURRENT APPLICATION NUMBER: US/09/962,832
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,077
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,280
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 259
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 113
; LENGTH: 2347
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-832-113

Query Match 38.7%; Score 465.8; DB 10; Length 2347;
Best Local Similarity 64.2%; Pred. No. 8.1e-126;
Matches 752; Conservative 2; Mismatches 399; Indels 18; Gaps 3;
Qy 14 GCGGGGGGGCGGGCGCTCTAGCGCTGTGTGCTGGCTCGCGCTGCGGGCGGGCG 73
Db 1 GCGGGGGGGCGATGGCTCGGTGGCTGGGGGCTGTGTGTGCTGGGAGCGCG 60
Qy 74 GCGCCCAATACGAAGCTACAGCTTCGAGCTTCCACGGGACGAGCTGATCGCCCTCG 133
Db 61 GGGCGCAGTACGAGAAGTACAGCTTCGCGGGCTTCCGCGGAGGACCTGATGCGCTGG 120
Qy 134 AGTCGGCTACCGCGCGCTGACAGTACAGCGCGGACGCTGGCGGAGGCTG 193
Db 121 CCGGGGCTACGGGACGCTCTGAGCAGTACAGGAGAGAGCTGGCGGAGCGCG 180
Qy 194 GCTACCTGGAGATCAGCTTCGCTGCGCTGCTGTGTGCGGACAGCGAGGCTTCTGCC 253
Db 181 GCTACCTGGAGGCGGCTGCGCTGCGCTGCTGTGTGCGGACAGCGAGGCTTCTGCC 240
Qy 254 ACCGCACTCAGG-----CCGCGCGGACGCGCGCGCGCGCGCTCGCA 304
Db 241 ACGCAACTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
Qy 305 GCTATCC-----CGAGCTCGGCTCTTCGGGGGCTGCTGCGCGCGCGCTCA 358
Db 301 ACGAGTGGGCTGCGAGCTGGGCTCTTGGCGCGCTCTTGGAGCGAGCGCGCTGCTGC 360

Qy 359 AGCGCTCAAGCAGGGCGCTGCGAGCTTCCGCCAGTCCCGAGCCAGCGCGAGTGTGG 418
Db 361 GCGCGTCAAGCGGACGCTGCGCGCTTCCAGGTGCGCTTACCCGCGCGAGTGTGC 420
Qy 419 CGGACTTCCAGCGCGCGAGCCCTACAAGTTCTTGCAGTTCTTCTTCAAGCAATA 478
Db 421 GTGACTTCCAGCGCGCGCTGCGCTTACAGTACCTGCTGCTGCTGCTTCAAGCTAAC 480
Qy 479 ATCTCCCAAGACCATCGCGCTGCTCACACTTCTTCTGAGCATCTCTGATGACGAA 538
Db 481 GGTGGAAGCGCGGTGCGCGCGCTTACACTTCTTCTTCCAGAGGAACCCGAGCAGC 540
Qy 539 TGATGAAGAGAAATGSCATATTATAA---GAGCCTGCTGCTGCGCGAGGACTACATTA 595
Db 541 TGACCGCAAGTATCTCACTTACTATCAGGGATGCTGGAGCTGCGCGAGTCCCTCA 600
Qy 596 AAGACTTGGAAACCAAGTCATATGAAGCTCTTTCATCCGAGCAGTGGGCGCATACAAG 655
Db 601 CGGACTAGAGGCGCGCTTACGAGCGCTGTTCTTCCGCGCTGTGAAGCTTCAACA 660
Qy 656 GTGGAAGCTGGAGAACATCCATCACAGCATGAGCTGGCGCTTCCGAGCTTCTTCAAG 715
Db 661 GCGGGATTTCCGAGCAGACAGGAGGACATGGAGCGGCTTGTCAAGTACCTGCGAG 720
Qy 716 CTTTACAGTGTCTCGAGCTCGAGGCTTCCAGGAGATCAAGGACTTCAAGGATT 775
Db 721 TCTTTCGCGCTGCTGCGCGCTGTGAAGGGGCGCATGAGCAGTGGACTTCAAGGACT 780
Qy 776 TCTACCTTCCATAGCAGATCATTTAGTGTGAGAAATTTGTGCTGCTGCTGCTGCTG 835
Db 781 TCTACCGCGCATAGCAGATCTTTGCGAGTCCCTGCGAGTGAAGTGGAGTGTGAGG 840
Qy 836 AGAAGCTCACCCAGTTTATAGAGGCTATCCGTTGAGAAATTTGTGCTGCTGCTGCTG 895
Db 841 CCAATTTGACCCCAATGTTGGTGGCTTCTTCTGCGGACAAAGTTGCTGCGCGCGCTGCTG 900
Qy 896 ATTACTTGCAGTTTGCCTTATTAAGTTGAAGCTTGAAGATGAGCGCGCTGCTGCTG 955
Db 901 ACTACCTGCACTTTCCTTACTTAAAGTTGAATGATGTCGCGCGAGCTGCGCGAGCGCG 960
Qy 956 TCAGCTATCTCTTTGATCAGAAATGACAGGTCATGCGAGCAGAACTGCTGCTGCTTACC 1015
Db 961 CCAGCTATCTCTTTCGACCCCAAGCAGCGCTATGCGAGCAGAACTGCTGCTGCTTACC 1020
Qy 1016 AGTACCAGGACACTTGGGCGCTCTCGATGAGCACTTCCAGCCAGACCTGGAAGCAG 1075
Db 1021 GGTTCACCGGCTCGTGGGCGCTTGAAGAGGAGGACTTCCAGCCCGCGAGGAGCCA 1080
Qy 1076 TTCAGTTCTTTAATGTGACCACTCCAGAGGAGCTGTATGCTTGTCAAGGAAATA 1135
Db 1081 TGCTTACCACACAGACCGCGAGCTGCGGAGCTGCTGGAGTTCACCCATGCTACC 1140
Qy 1136 TAATGATGATGATGAGGAGAAAGTTGTGA 1166
Db 1141 TGCAGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1171

RESULT 5

US-10-439-388-64
; Sequence 64, Application US/10439388
; Publication No. US20030228617A1
; GENERAL INFORMATION:
; APPLICANT: Aune, Thomas M
; APPLICANT: Olsen, Nancy J
; TITLE OF INVENTION: Method for Predicting Autoimmune Disease
; FILE REFERENCE: 1242/68
; CURRENT APPLICATION NUMBER: US/10/439,388
; CURRENT FILING DATE: 2003-05-16
; PRIOR APPLICATION NUMBER: US 60/381,055
; PRIOR FILING DATE: 2002-05-16
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.2


```
RESULT 7
US-10-045-815-3
; Sequence 3, Application US/10045815
; Publication No. US20020160498A1
; GENERAL INFORMATION:
; APPLICANT: Wadhwah, Renu
; APPLICANT: Sugihara, Takashi
; APPLICANT: Ohide, Akiko
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; FILE REFERENCE: 06501-091001
; CURRENT APPLICATION NUMBER: US/10/045,815
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: PCT/JP00/02731
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: JP 11/118806
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2600
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (52)...(2259)
US-10-045-815-3

Query Match      12.9%; Score 154.8; DB 14; Length 2600;
Best Local Similarity 53.4%; Pred. No. 1.7e-34;
Matches 436; Conservative 0; Mismatches 347; Indels 33; Gaps 4;

Qy 222 CGCTTGTGCGGACGAGGCGCTTGTGCACCGCACTGCAGCGCGCGCGCGAGCC 281
Db |||||
Qy 276 CGCACCACAGTGTGCGGCGCTTCCGTGGAGCTGACCGCGAGCTGTGCCCGAGCCC 335
Db |||||
Qy 282 CGAGCCCGCGGCGCTTCCAGCGCGCGAGCGCTTACAGTTCCTGCGAGTTCGC 341
Db |||||
Qy 447 CAGCGAGAGATGGAGCTGGAGTTCGCGAAGCGAGCGCCCTACACTACCTGCGAGTGC 506
Db |||||
Qy 462 TTACTTTCAAGGCAATATCTCCCAAGCCATCGCCGCTGCTCACCTTTCTACTGAA 521
Db |||||
Qy 507 CTACTTCAAGATCAACAAGTTGGAGAAAGCTGTTGCTGCAGCACACACCTTCTTCGTGG 566
Db |||||
Qy 522 GCATCCTGATGCGAAATGATGAAGAGAAACATGGCATATTATAGAGCGCTGCTGGTG- 580
Db |||||
Qy 567 CAATCTGAGCATGGAATGACAGACACCTAGACTATTACCAACCACTGTGTGAGT 626
Db |||||
Qy 581 --CGAGGACTACATAAGAGCTCGAAACCAAGTTCATATGAAGCGCTGTTTATCCGAGC 638
Db |||||
Qy 627 GAAGGAGCGGAGCTTCAAGGATCTTGAGACTCAACCCCATATGCAAGAATTTGCACTGG 686
Db |||||
Qy 639 AGTGGGGGATACACGGGTGAGACTGGAGAACATCCATCAGACATGGAGCTGGCCCT 698
Db |||||
Qy 687 AGTGGACTCTACTCAGAGGAACAGCCACAGGAAGCTGTGCCCACTTAGAGGGCGGCT 746
Db |||||
Qy 699 TCCGAGCTTCTTCAAGCGCTTTTACGAGTGTCTGAGCGCTGCGAGGGTTCAGGGAGAT 758
Db |||||
Qy 747 GCAAGAATACTTTGTGGCTATGAGGAGTGGCGTCCCTCTGCGAAGGCGCTTATGACTA 806
Db |||||
Qy 759 CAAGGACTTCA-----GGATTTCTACCTTTCCATAGCAGATCATTA 800
Db |||||
Qy 807 CGATGGCTACAACTACCTTGTAGTCAACGCTGACCTCTTCCAGGCGCATCAGATCATTA 866
Db |||||
Qy 801 TGTAAGAGTCTGGAATGCAAAATACAGTGTGAAGAGAACCT---CACCCCGATTATAGG 857
Db |||||

RESULT 8
US-10-045-815-1
; Sequence 1, Application US/10045815
; Publication No. US20020160498A1
; GENERAL INFORMATION:
; APPLICANT: Wadhwah, Renu
; APPLICANT: Sugihara, Takashi
; APPLICANT: Ohide, Akiko
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; FILE REFERENCE: 06501-091001
; CURRENT APPLICATION NUMBER: US/10/045,815
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: PCT/JP00/02731
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: JP 11/118806
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2829
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (52)...(1140)
US-10-045-815-1

Query Match      12.9%; Score 154.8; DB 14; Length 2829;
Best Local Similarity 53.4%; Pred. No. 1.7e-34;
Matches 436; Conservative 0; Mismatches 347; Indels 33; Gaps 4;

Qy 222 CGCTTGTGCGGACGAGGCGCTTGTGCCACCGCACTGCAGCGCGCGCGCGAGCC 281
Db |||||
Qy 276 CGCACCACAGTGTGCGGCGCTTCCGTGGAGCTGACCGCGAGCTGTGCCCGAGCCC 335
Db |||||
Qy 282 CGAGCCCGCGGCGCTTCCAGCGCGCGAGCGCTTACAGTTCCTGCGAGTTCGC 341
Db |||||
Qy 336 GGCCAGGCGCTCGGGCGCGCGCGCTGCGACCTGACCTTCTTCGGGGCGCTTTCGG 395
Db |||||
Qy 342 CGCGCGGACCTCCTCAAGCGCTGCAAGCGGCGCTGCGAGCGCTTCCCGAGTCCAGCC 401
Db |||||
Qy 396 TCGCGCTGCTGCTGCGCGCGCTGCTGCGGCGCGCGCGCGCGCGCGCGCGCTC-----GCT 446
Db |||||
Qy 402 CAGCCCGGAGGTGCTGGCGGACTTCCAGCGCGCGAGCGCTTACAGTTCCTGCGAGTTCGC 461
Db |||||
Qy 447 CAGCGAGAGATGGAGCTGGAGTTCGCGAAGCGAGCGCCCTACACTACCTGCGAGTGC 506
Db |||||
Qy 462 TTACTTTCAAGGCAATATCTCCCAAGCCATCGCCGCTGCTCACCTTTCTACTGAA 521
Db |||||
Qy 507 CTACTTCAAGATCAACAAGTTGGAGAAAGCTGTTGCTGCAGCACACACCTTCTTCGTGG 566
Db |||||
Qy 522 GCATCCTGATGCGAAATGATGAAGAGAAACATGGCATATTATAGAGCGCTGCTGGTG- 580
Db |||||
Qy 567 CAATCTGAGCATGGAATGACAGACACCTAGACTATTACCAACCACTGTGTGAGT 626
Db |||||
Qy 581 --CGAGGACTACATAAGAGCTCGAAACCAAGTTCATATGAAGCGCTGTTTATCCGAGC 638
Db |||||
Qy 627 GAAGGAGCGGAGCTTCAAGGATCTTGAGACTCAACCCCATATGCAAGAATTTGCACTGG 686
Db |||||
Qy 639 AGTGGGGGATACACGGGTGAGACTGGAGAACATCCATCAGACATGGAGCTGGCCCT 698
Db |||||
Qy 687 AGTGGACTCTACTCAGAGGAACAGCCACAGGAAGCTGTGCCCACTTAGAGGGCGGCT 746
Db |||||
Qy 699 TCCGAGCTTCTTCAAGCGCTTTTACGAGTGTCTGAGCGCTGCGAGGGTTCAGGGAGAT 758
Db |||||
Qy 747 GCAAGAATACTTTGTGGCTATGAGGAGTGGCGTCCCTCTGCGAAGGCGCTTATGACTA 806
Db |||||
Qy 759 CAAGGACTTCA-----GGATTTCTACCTTTCCATAGCAGATCATTA 800
Db |||||
Qy 807 CGATGGCTACAACTACCTTGTAGTCAACGCTGACCTCTTCCAGGCGCATCAGATCATTA 866
Db |||||
Qy 801 TGTAAGAGTCTGGAATGCAAAATACAGTGTGAAGAGAACCT---CACCCCGATTATAGG 857
Db |||||
```

Db 627 GAAGAGCCGACCTTCAAGAGTCTTGAGACTCAACCCCATATGCAAGAAATTCGACTGG 686
QY 639 AGTGGGCGCATACACCGGTGAGAACTGGAGAACTCATCACAGACATGAGCTGGCCCT 698
Db 687 AGTGGGACTCTACTCAGAGGAACAGCCACAGGAAGCTGTGCCACCTAGAGCGGCGCT 746
QY 699 TCCGAGCTTCTTCAAGCCTTTAGGAGTGTCTCGAGCCTCGGAGGTTCCAGGGAGAT 758
Db 747 GCAAGAACTACTTTGTGGCCTATGAGGAGTGGCGCTCTGCGAAGGGCCCTATGACTA 806
QY 759 CAAGGACTTCAA-----CGATTTCTACCTTTCCATAGCAGATCATTA 800
Db 807 CGATGGCTACACTACCTTGTAGTACACGCTGACTCTTCCAGGCCATCAGATCATTA 866
QY 801 TGTAGAATTTCTGAATGCAAAATACAGTGTGAAGAAAGCT---CACCCAGTTATAGG 857
Db 867 CATCAGTCTCCTAACTGTAAAGACAGTGTGTACGAGGCTTGTCTTCCCAACCAAGTCG 926
QY 858 AGGCTATCCGGTTTCAGAAATTTGGGTACCATGTATCATCTACTGAGTTTCCTTATTA 917
Db 927 AGAAGGCCCTTTGAAGACTTCTCCCATGCGATATTAATATCTGCAGTTTGCCTACTA 986
QY 918 TAAGTTGAACACCTGGAAGATGAGCCCTGTCAGTCTGAGTATCTGCTCTTTGATCA 977
Db 987 TAACATTTGGGAATATATACAAAGCTGGTGAATGTGCCAAGACCTATCTTCTTCTTCCC 1046
QY 978 GAATGCAAGGTCTAGCAGCAAGACCTGGTGTATTA 1013
Db 1047 CAATGACGAGGTGATGAACCAAAATTTGGCCTATTA 1082

RESULT 9

US-10-027-632-161474/c

; Sequence 161474, Application US/10027632

; Publication No. US20030204075A9

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027.632

; CURRENT FILING DATE: 2002-04-30

; PRIOR APPLICATION NUMBER: US 60/218,006

; PRIOR FILING DATE: 2000-07-12

; PRIOR APPLICATION NUMBER: US 60/198,676

; PRIOR FILING DATE: 2000-04-20

; PRIOR APPLICATION NUMBER: US 60/193,483

; PRIOR FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: US 60/185,218

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/167,363

; PRIOR FILING DATE: 1999-11-23

; PRIOR APPLICATION NUMBER: US 60/156,358

; PRIOR FILING DATE: 1999-09-28

; PRIOR APPLICATION NUMBER: US 60/146,002

; PRIOR FILING DATE: 1999-08-09

; NUMBER OF SEQ ID NOS: 325720

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 161474

; LENGTH: 892

; TYPE: DNA

; ORGANISM: Human

US-10-027-632-161474

Query Match

Best Local Similarity 12.8%; Score 153.8; DB 13; Length 892;

Matches 155; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 463 TACTTCAAGGCAATAATCTCCCAAGCCATCGCCGCTGCTCACACCTTTCTACTGAAG 522
Db 228 TCCTTTATAGGCAATAATCTCCCAAGCCATCGCCGCTGCTCACACCTTTCTACTGAAG 169
QY 523 CATCTCTGATGACGAATGATGAAGAGGAACATGGCATATTATAGAGCCTGCTGCTGCC 582

Db 168 CATCTCTGATGATGAATGATGAAGAGGAACATGGCATATTATAGAGCCTGCTGCTGCC 109
QY 583 GAGGACTACATTAAAGACCTGGARACCAAGTCATATGAA 621
Db 108 GAGGACTACATTAAAGACCTGGARACCAAGTCATATGAA 70

RESULT 10

US-10-027-632-161475/c

; Sequence 161475, Application US/10027632

; Publication No. US20030204075A9

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027.632

; CURRENT FILING DATE: 2002-04-30

; PRIOR APPLICATION NUMBER: US 60/218,006

; PRIOR FILING DATE: 2000-07-12

; PRIOR APPLICATION NUMBER: US 60/198,676

; PRIOR FILING DATE: 2000-04-20

; PRIOR APPLICATION NUMBER: US 60/193,483

; PRIOR FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: US 60/185,218

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/167,363

; PRIOR FILING DATE: 1999-11-23

; PRIOR APPLICATION NUMBER: US 60/156,358

; PRIOR FILING DATE: 1999-09-28

; PRIOR APPLICATION NUMBER: US 60/146,002

; PRIOR FILING DATE: 1999-08-09

; NUMBER OF SEQ ID NOS: 325720

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 161475

; LENGTH: 892

; TYPE: DNA

; ORGANISM: Human

US-10-027-632-161475

Query Match

Best Local Similarity 12.8%; Score 153.8; DB 13; Length 892;

Matches 155; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 463 TACTTCAAGGCAATAATCTCCCAAGCCATCGCCGCTGCTCACACCTTTCTACTGAAG 522
Db 228 TCCTTTATAGGCAATAATCTCCCAAGCCATCGCCGCTGCTCACACCTTTCTACTGAAG 169
QY 523 CATCTCTGATGACGAATGATGAAGAGGAACATGGCATATTATAGAGCCTGCTGCTGCC 582
Db 168 CATCTCTGATGATGAATGATGAAGAGGAACATGGCATATTATAGAGCCTGCTGCTGCC 109
QY 583 GAGGACTACATTAAAGACCTGGARACCAAGTCATATGAA 621
Db 108 GAGGACTACATTAAAGACCTGGARACCAAGTCATATGAA 70

RESULT 11

US-10-027-632-161476/c

; Sequence 161476, Application US/10027632

; Publication No. US20030204075A9

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027.632

; CURRENT FILING DATE: 2002-04-30

; PRIOR APPLICATION NUMBER: US 60/218,006

; PRIOR FILING DATE: 2000-07-12

; PRIOR APPLICATION NUMBER: US 60/198,676

; PRIOR FILING DATE: 2000-04-20

;; PRIOR APPLICATION NUMBER: US 60/193,483
;; PRIOR FILING DATE: 2000-03-29
;; PRIOR APPLICATION NUMBER: US 60/185,218
;; PRIOR FILING DATE: 2000-02-24
;; PRIOR APPLICATION NUMBER: US 60/167,363
;; PRIOR FILING DATE: 1999-11-23
;; PRIOR APPLICATION NUMBER: US 60/156,358
;; PRIOR FILING DATE: 1999-09-28
;; PRIOR APPLICATION NUMBER: US 60/146,002
;; PRIOR FILING DATE: 1999-08-09
;; NUMBER OF SEQ ID NOS: 325720
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 161476
;; LENGTH: 892
;; TYPE: DNA
;; ORGANISM: Human
US-10-027-632-161476

Query Match 12.8%; Score 153.8; DB 13; Length 892;
Best Local Similarity 97.5%; Pred. No. 1.9e-34;
Matches 155; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 463 TACTTCAAGGCAAAATAATCTCCCAAGCCATCGCGCTGCTCACACCTTTCTACTGAAG 522
Db 228 TCCTTTTAGGCAAAATAATCTCCCAAGCCATCGCGCTGCTCACACCTTTCTACTGAAG 169

QY 523 CATCTGTATGACGAATGATGAAGAGAACATGGCATATTATAGAGCCTCGCTGTGCC 582
Db 168 CATCTGTATGAYGAATGATGAAGAGAACATGGCATATTATAGAGCCTCGCTGTGCC 109

QY 583 GAGGACTACATTAAAGACCTGGAAACCAAGTCATATGAA 621
Db 108 GAGGACTACATTAAAGACCTGGAAACCAAGTCATATGAA 70

RESULT 12
US-10-027-632-161474/c
;; Sequence 161474, Application US/10027632
;; GENERAL INFORMATION:
;; APPLICANT: Wang, David G.
;; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;; FILE REFERENCE: 108827.129
;; CURRENT APPLICATION NUMBER: US 60/193,483
;; PRIOR FILING DATE: 2000-04-30
;; PRIOR APPLICATION NUMBER: US 60/218,006
;; PRIOR FILING DATE: 2000-07-12
;; PRIOR APPLICATION NUMBER: US 60/198,676
;; PRIOR FILING DATE: 2000-04-20
;; PRIOR APPLICATION NUMBER: US 60/193,483
;; PRIOR FILING DATE: 2000-03-29
;; PRIOR APPLICATION NUMBER: US 60/185,218
;; PRIOR FILING DATE: 2000-02-24
;; PRIOR APPLICATION NUMBER: US 60/167,363
;; PRIOR FILING DATE: 1999-11-23
;; PRIOR APPLICATION NUMBER: US 60/156,358
;; PRIOR FILING DATE: 1999-09-28
;; PRIOR APPLICATION NUMBER: US 60/146,002
;; PRIOR FILING DATE: 1999-08-09
;; NUMBER OF SEQ ID NOS: 325720
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 161474
;; LENGTH: 892
;; TYPE: DNA
;; ORGANISM: Human
US-10-027-632-161474

Query Match 12.8%; Score 153.8; DB 14; Length 892;
Best Local Similarity 97.5%; Pred. No. 1.9e-34;
Matches 155; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 463 TACTTCAAGGCAAAATAATCTCCCAAGCCATCGCGCTGCTCACACCTTTCTACTGAAG 522

QY 463 TACTTCAAGGCAAAATAATCTCCCAAGCCATCGCGCTGCTCACACCTTTCTACTGAAG 522

Db 228 TCCTTTTAGGCAAAATAATCTCCCAAGCCATCGCGCTGCTCACACCTTTCTACTGAAG 169

QY 523 CATCTGTATGACGAATGATGAAGAGAACATGGCATATTATAGAGCCTCGCTGTGCC 582

Db 168 CATCTGTATGAYGAATGATGAAGAGAACATGGCATATTATAGAGCCTCGCTGTGCC 109

QY 583 GAGGACTACATTAAAGACCTGGAAACCAAGTCATATGAA 621

Db 108 GAGGACTACATTAAAGACCTGGAAACCAAGTCATATGAA 70

RESULT 13
US-10-027-632-161475/c
;; Sequence 161475, Application US/10027632
;; GENERAL INFORMATION:
;; APPLICANT: Wang, David G.
;; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;; FILE REFERENCE: 108827.129
;; CURRENT APPLICATION NUMBER: US 60/218,006
;; PRIOR FILING DATE: 2000-07-12
;; PRIOR APPLICATION NUMBER: US 60/198,676
;; PRIOR FILING DATE: 2000-04-20
;; PRIOR APPLICATION NUMBER: US 60/193,483
;; PRIOR FILING DATE: 2000-03-29
;; PRIOR APPLICATION NUMBER: US 60/185,218
;; PRIOR FILING DATE: 2000-02-24
;; PRIOR APPLICATION NUMBER: US 60/167,363
;; PRIOR FILING DATE: 1999-11-23
;; PRIOR APPLICATION NUMBER: US 60/156,358
;; PRIOR FILING DATE: 1999-09-28
;; PRIOR APPLICATION NUMBER: US 60/146,002
;; PRIOR FILING DATE: 1999-08-09
;; NUMBER OF SEQ ID NOS: 325720
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 161475
;; LENGTH: 892
;; TYPE: DNA
;; ORGANISM: Human
US-10-027-632-161475

Query Match 12.8%; Score 153.8; DB 14; Length 892;
Best Local Similarity 97.5%; Pred. No. 1.9e-34;
Matches 155; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 463 TACTTCAAGGCAAAATAATCTCCCAAGCCATCGCGCTGCTCACACCTTTCTACTGAAG 522

Db 228 TCCTTTTAGGCAAAATAATCTCCCAAGCCATCGCGCTGCTCACACCTTTCTACTGAAG 169

QY 523 CATCTGTATGACGAATGATGAAGAGAACATGGCATATTATAGAGCCTCGCTGTGCC 582

Db 168 CATCTGTATGAYGAATGATGAAGAGAACATGGCATATTATAGAGCCTCGCTGTGCC 109

QY 583 GAGGACTACATTAAAGACCTGGAAACCAAGTCATATGAA 621

Db 108 GAGGACTACATTAAAGACCTGGAAACCAAGTCATATGAA 70

RESULT 14
US-10-027-632-161476/c
;; Sequence 161476, Application US/10027632
;; GENERAL INFORMATION:
;; APPLICANT: Wang, David G.
;; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;; FILE REFERENCE: 108827.129
;; CURRENT APPLICATION NUMBER: US 60/218,006
;; PRIOR FILING DATE: 2000-07-12
;; PRIOR APPLICATION NUMBER: US 60/198,676

QY 463 TACTTCAAGGCAAAATAATCTCCCAAGCCATCGCGCTGCTCACACCTTTCTACTGAAG 522

Db 228 TCCTTTTAGGCAAAATAATCTCCCAAGCCATCGCGCTGCTCACACCTTTCTACTGAAG 169

QY 523 CATCTGTATGACGAATGATGAAGAGAACATGGCATATTATAGAGCCTCGCTGTGCC 582

Db 168 CATCTGTATGAYGAATGATGAAGAGAACATGGCATATTATAGAGCCTCGCTGTGCC 109

QY 583 GAGGACTACATTAAAGACCTGGAAACCAAGTCATATGAA 621

Db 108 GAGGACTACATTAAAGACCTGGAAACCAAGTCATATGAA 70

288	CGGACCCGACGCTGAGCCAGGACCCGGGCGCGCGCGCTGCAGACCTGGCTTCTTC	347
Db		
328	GGGGCCCTGCTGCGCGCGCGCACTGCCTCAAGCGCTGCAAGCAGGGCTGCCAGCCTTC	387
QY		
348	GGAGCCGTGCTGCGCGCTGCGGCTGCTAGCGCGCTGCTGGGCGCGCCCTCTGCCAC	407
Db		
388	CGCCAGTCCGAGCCAGCCGCGAGGCTGCTGGGGACTTCCAGCGCGCGAGCCCTACAAG	447
QY		
408	TTGCTG-----AGTGAAGAACTGGACCTGGAGTTCAAAGCGGAGCCGTACAAC	458
Db		
448	TTCTGTCAGTTGCTTACTTTCAAGGGAATAATCTCCCCAAGAGCCATGCCGCTGCTAC	507
QY		
459	TACCTGAGGTGCGCTATTTCAGAGTAACAAGCTGGAGAAAGCTGTGCTCGGCACAC	518
Db		
508	ACCTTTTACTGAAGCATCCTGATGACGAAATGATGAAGAGGAACATGGCATATTATAAG	567
QY		
519	ACCTTCTTTGTGGCAATCCTGAGCACATGGAGATGCGGCAGAACTCGACTATTACCA	578
Db		
568	AGCTGCTGCTGTG--CCGAGGACTACATTABAGACCTGGAAACCAAGTCATATGAAGC	624
QY		
579	ACCATGTCTGGGTGAAGGAGGAGACTTCAGGGATCTCGAGGCCAAGCCCAATGCGAT	638
Db		
625	CTGTTTCATCCGACGATGCGGGCATACAACGTTGAGAACTGGAGAACATCCATCACAGC	684
QY		
639	GAGTTTCGGCTGGGGTACGACTCTACTCAGAGGAGAGCCACAGAGAACTGTGCCCCAC	698
Db		
685	ATGAGAGTGGCCCTTCCGACTCTTCAAAGCCTTTTACGAGTGCTCGAGCCTGGAG	744
QY		
699	CTGAGCGGCACTGCAAGAGTACTTTGTGGCGGATGAGGAGTGCCTGCGCCTCTGGAA	758
Db		
745	GGTTCACGGAGATCAAGGACTTCAA-----GGATTTCTACCTTTCC	786
QY		
759	GGGCGCTATGACTACGACGGCTACAACTACTTAGACTACAGGCGTGACCTCTTCAGGCC	818
Db		
787	ATACAGATCATTTATGTAGAGTTCTGGAATGCAAAATACAGTGTGAGAGAACT---C	843
QY		
819	ATCACAGATCATTTAGTTCAGGTCTCAACTGTAGCAGAACTGTGTCCAGGAGCTGGCT	878
Db		
844	ACCCAGTTATAGAGGCTATCCGGTTGAGAAATTTGTGGTACCATGTATCATTTACTTG	903
QY		
879	TCCACCCAAAGTAGGAAAAGCCCTTGAAGACTTCTCCCTTCACACTATAATTACCTA	938
Db		
904	CAGTTTCCCTATTATAGTTGAACGACTGAGAGATGAGCCCGCTGTGCACTAGTAT	963
QY		
939	CAGTTTCCCTACTACACACTTGGAACTATACAAAGCTATTGAATGTGCCAAGACCTAC	998
Db		
964	CTGCTCTTTGATCAGAATGACAAGSTCATGCAAGAGAACCTGGTGTTATAC	1014
QY		
999	CTCCTCTTCTTTCCCAATGATCAGGTGTATGACCAAGAACTGGCTATTATA	1049
Db		

Search completed: January 30, 2004, 19:07:31
Job time : 410.168 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2004, 08:05:34 ; Search time 67.8238 Seconds
(without alignments)
7828.876 Million cell updates/sec

Title: US-09-729-674-1_COPY_63_1265
Perfect score: 1203
Sequence: 1 atggagccggggccggggg.....tggaaactggaggagaccagc 1203

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	126	10.5	618	3	US-09-328-111-215
2	64	5.3	846	4	US-09-352-991A-1177
C 3	64	5.3	1083	4	US-09-352-991A-1387
4	64	5.3	1089	4	US-09-352-991A-1221
C 5	60.4	5.0	30001	1	US-08-125-468-1
C 6	60.4	5.0	30001	2	US-08-474-933-1
7	59	4.9	2481	1	US-08-324-243-35
8	59	4.9	2481	1	US-08-532-390-35
9	59	4.9	2481	3	US-08-717-294-35
10	59	4.9	2481	5	PCT-US95-11511-35
11	59	4.9	2571	4	US-09-552-950-4
C 12	55.4	4.6	1140	3	US-09-105-537-15
C 13	55.4	4.6	13613	3	US-09-105-537-3
14	54.4	4.5	3292	3	US-09-320-878-22
15	54.4	4.5	3292	4	US-09-141-908-15
16	54.4	4.5	3292	4	US-09-557-440-22
17	52.6	4.4	1686	4	US-09-352-991A-6509
C 18	52.6	4.4	1740	4	US-09-352-991A-6553
C 19	52.6	4.4	1899	4	US-09-352-991A-6558
20	52.6	4.4	2469	4	US-09-252-991A-6519
21	52	4.3	1155	4	US-09-252-991A-15330
22	52	4.3	3624	1	US-07-951-715A-6
23	52	4.3	3624	2	US-08-459-448A-6
24	52	4.3	3624	3	US-08-459-595A-6
25	52	4.3	3624	3	US-08-459-504B-6
26	52	4.3	3624	3	US-08-459-444-6
27	52	4.3	3624	3	US-09-053-549-7

C 28	52	4.3	3624	4	US-09-547-422-6	Sequence 6, Appli
29	51.6	4.3	1419	4	US-09-252-991A-11636	Sequence 11636, A
C 30	51.6	4.3	1587	4	US-09-252-991A-11720	Sequence 11720, A
31	51.6	4.3	2304	4	US-09-252-991A-11802	Sequence 11802, A
32	50.6	4.2	942	4	US-09-252-991A-3976	Sequence 3976, Ap
33	50.6	4.2	1521	4	US-09-252-991A-4046	Sequence 4046, Ap
C 34	50.4	4.2	5228	4	US-09-428-711A-15	Sequence 15, Appl
35	50.4	4.2	2928	4	US-09-252-991A-12178	Sequence 12178, A
C 36	50.4	4.2	3678	4	US-09-126-280-3	Sequence 12417, A
37	50.2	4.2	2635	3	US-09-126-280-3	Sequence 3, Appli
38	50.2	4.2	2670	3	US-09-126-280-1	Sequence 1, Appli
39	50.2	4.2	2961	2	US-08-407-875-1	Sequence 1, Appli
40	50.2	4.2	2961	4	US-09-277-858-1	Sequence 1, Appli
41	50	4.2	681	4	US-09-252-991A-7273	Sequence 7273, Ap
C 42	50	4.2	2541	4	US-09-252-991A-7209	Sequence 7209, Ap
C 43	50	4.2	3099	4	US-09-252-991A-7487	Sequence 7487, Ap
44	49.8	4.1	13842	3	US-09-105-537-30	Sequence 30, Appl
45	49.8	4.1	36778	3	US-09-105-537-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-09-328-111-215/C
; Sequence 215, Application US/09328111
; Patent No. 6262333
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dertli, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/328,111
; CURRENT FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: US 60/088,801
; EARLIER FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 215
; LENGTH: 618
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: misc feature
; LOCATION: (1)...(618)
; OTHER INFORMATION: n = A,T,C or G
US-09-328-111-215

Query Match	10.5%	Score 126;	DB 3;	Length 618;
Best Local Similarity	85.9%	Pred. No. 5.4e-21;		
Matches 165;	Conservative 0;	Mismatches 24;	Indels 3;	Gaps 3;
QY	1014	CCAGTACACAGGACACTTGGGCGCTTCGGATGACACTCCAGCCAGACTGAAGC	1073	
DB	534	CNAGTCCNNAGNCCANTGGGCGCTTNGATGAGCANTCCAGCCAGNCTTAANC	475	
QY	1074	AGTTCAGTCTTTTAATGTGACCA-CACTCCAGAGGAGC-TGTATGACTTT-CTTAAGGA	1130	
DB	474	AGTTCAGTCTTTAAATGTGNCACCACTCCAGAGGAGCTTGTATGACTTTGNTAAGGA	415	
QY	1131	AAATATATGATGATGATGAGGAGAGTGTGGATATGATGATGACTCTTGAAC	1190	

Db 414 AATATTAATGATGATGATGAGGAGAAAGTTGTGGAATATGTGGATGACCTCTTGGAACT 355

Qy 1191 GGAGGAGACGAC 1202
|||||

Db 354 GGAGGAGACGAC 343

RESULT 2

US-09-252-991A-1177
; Sequence 1177, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.1136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1177
; LENGTH: 846
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1177

Query Match 5.3%; Score 64; DB 4; Length 846;
Best Local Similarity 46.6%; Pred. No. 3.4e-06;
Matches 205; Conservative 0; Mismatches 235; Indels 0; Gaps 0;

Qy 12 GCGCGCGGGGCGCGCGCGCTGTAGCGCTGTGTGCGTGTGCGCTGTGCGCGCTGTGCGCGCGG 71
Db 373 GCTCGCGGTGACCCGGCGGATTCCTGCGCTGTCTGCGCGTGTGCGCGCGCGCGCGCTGTGT 432

Qy 72 GCGCGGCCAATACGAAGCTACAGCTTCCGAGCTTCCACGGGACGAGCTGATGCCGCT 131
Db 433 GAACGTCACCTCGGGCTCGGCTCGGTGCGCATTCCTACCTGCGGCGCTACTCCGCGCG 492

Qy 132 CGAGTCGGCTACCGGACGCGCTGGACAGTACAGCGGCGAGACTTGGCGCGGAGAGCST 191
Db 493 GCAGTTCGCAAGAGAGGAGTGAAGCAGCGCCCTGCGCGCGAGCTGTGCGGCCATGGCAT 552

Qy 192 KGGCTACTTGAGATCAGCTTGGCTGCGCTGACCGCTTGTCTGCGGACAGCGAGCGCTTCTG 251
Db 553 CCAGGTCCTGGTGTGAGCCCGGGCGATCTGGACGCCATCTGGGGCAGATGCCAG 612

Qy 252 CCACCGCAATGACAGCGCGCGCGAGCCGAGCGCGCGCGCTTCCAGAGCTATCC 311
Db 613 CGAGGGCGAGCGCGCCCTGGCGCAGCGCCCGAGCGCGCTGCGCGACCTCTATCGCGATAC 672

Qy 312 CGAGCTCGGCTCTTTCGGGGCGCTGTGCGCGCGGCGACTGTCTCAAGCGCTGCAAGCA 371
Db 673 CTACCTCGGTTCTTCAGGCCAAGAGAGGCGCGCGCAACAGCGCGACCAAGCGCGC 732

Qy 372 GGGCTGCCAGCTTCCGCCAGTCCAGCCAGCGGAGGTGTGTGCGGAGCTTCCAGCG 431
Db 733 CGAGCTCGCGCGCGCGTGTGATCGCGCTCACCGCGCCAAAGCGCGGACCGCTACCG 792

Qy 432 CCGCGAGCCCTACAGTTCC 451
Db 793 GTGGGCGCGAGCTGCGCC 812

RESULT 3

US-09-252-991A-11387/c
; Sequence 1387, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

	TELEFAX: (201)831-3305	
:	INFORMATION FOR SEQ ID NO: 1:	
:	SEQUENCE CHARACTERISTICS:	
:	LENGTH: 30001 base pairs	
:	TYPE: nucleic acid	
:	STRANDEDNESS: single	
:	TOPOLOGY: linear	
:	MOLECULE TYPE: DNA (genomic)	
:	US-08-125-468-1	
	Query Match	5.0%; Score 60.4; DB 1; Length 30001;
	Best Local Similarity	47.9%; Pred. No. 8.6e-05;
	Matches 203; Conservative	1; Mismatches 217; Indels 3; Gaps 1;
QY	39 GCTGCTGTGGTGGCTGTGGCGCTGGCGCCGGCGCGGCCCAATACGAACGCCTACAGTT 98	
Db	6270 GTTGCTGCGCAGCGCACCGCCCAAATCGGCGTCGCCGTGGAATGGGCGTCGAACTCGC 6211	
QY	99 CCGCAGCTTCCCACGGGACGAGCTGATGCCGCTCGAGTCGGGCTACCGGCACGGCGTGA 158	
Db	6210 CGCGCTGCAGCCCGGSCCGACTCCGTGCTCGTCACTCCGCCACGCCGACGGGCGCAC 6151	
QY	159 CAAGTACACGGCGAGCATCTGGCCGAGAGSTKGGCTACTTGGAGATCAGCCTTCGGGCT 218	
Db	6150 CGAGGAGACC CGGGTGCCTTGGCTGGTCCGCGCCGACGCGCGCCGACGACCGTGCACA 6091	
QY	219 GCACGCGTTGCTCGCGACAGCGAGGCGCTTCTGCCACCGCAACTGCACGCGCGCGCCGCA 278	
Db	6090 GCAGCTCGGATTCGCGCTGATCGGCGAATCCACCGAGTCTGGCTCAACGCGGACGTCAC 6031	
QY	279 GCCGAGCCCGCCGCGGCGCTGCCAGCTATCCGAGTATCCGAGCTGGCGCTCTTCGGGGGCGCTGCT 338	
Db	6030 CCTCGACGCGGACCTCCGCGCGACAGCAACACCTGTGCACACCGAGCCCCGACCCCT 5971	
QY	339 GCGCGCGCGCACTCGCTCAAGCGTGAAGAGGGCGCTGCAGGCTTCCGCCAGTCCA 398	
Db	5970 GTGTGCTGTCGCTTCCCGGACCGCGCAAGTGGCGGTGGTGACACGGTGGACACCGA 5911	
QY	399 GCCCA3CGCCGAGGTGCTGGCGGACTTCCAGCGCGCGAGCGCTTACAGTTCTTCGCAAGTT 458	
Db	5910 ---CGGCGCGAGGACCGCGAGCTGGTCCGCGCGGCTGCCGACAGAATCGGGCGGG 5854	
QY	459 CGCT 462	
Db	5853 CGCT 5850	
 RESULT 6 US-08-474-933-1/c Sequence 1, Application US/08474933 Patent No. 5866410 GENERAL INFORMATION: APPLICANT: Ryan, Michael J. APPLICANT: Lotvin, Jason A. APPLICANT: Strathy, Nancy E. APPLICANT: Fantini, Susan E. TITLE OF INVENTION: Cloning of the biosynthetic pathway for TITLE OF INVENTION: chlorotetracycline and tetracycline Formation and cosmid TITLE OF INVENTION: useful therein NUMBER OF SEQUENCES: 1 CORRESPONDENCE ADDRESS: ADDRESS: American Cyanamid Company STREET: One Cyanamid Plaza CITY: Wayne STATE: New Jersey COUNTRY: USA ZIP: 07470 COMPUTER READABLE FORM: MEDIUM TYPE: Floppy disk COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS SOFTWARE: PatentIn Release #1.0, Version #1.25 CURRENT APPLICATION DATA:		

ORGANISM: Streptomyces venezuelae
S-09-105-537-15

Query Match 4.6%; Score 55.4; DB 3; Length 1140;
Best Local Similarity 46.3%; Pred. No. 0.00043;
Matches 179; Conservative 1; Mismatches 207; Indels 0

Db 3209 GCTGCGGTGATCGACGCCGCTGCGCGA 3235
|||||

RESULT 15

US-09-141-908-15
; Sequence 15, Application US/09141908
; Patent No. 6503741
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: Combinatorial Polyketide Libraries Produced Using a
; FILE OF INVENTION: Modular PKS Gene Cluster as Scaffold
; FILE REFERENCE: 300622002100
; CURRENT APPLICATION NUMBER: US/09/141,908
; CURRENT FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: CIP OF 09/073,538
; EARLIER FILING DATE: 1998-05-06
; EARLIER APPLICATION NUMBER: CIP OF 08/846,247
; EARLIER FILING DATE: 1997-04-30
; EARLIER APPLICATION NUMBER: PROV. 60/076,919
; EARLIER FILING DATE: 1998-03-05
; EARLIER APPLICATION NUMBER: PROV. 60/087,080
; EARLIER FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; TYPE: DNA
; LENGTH: 3292
; ORGANISM: Streptomyces venezuelae
US-09-141-908-15

Query Match 4.5%; Score 54.4; DB 4; Length 3292;
Best Local Similarity 46.0%; Pred. No. 0.0011;
Matches 178; Conservative 1; Mismatches 208; Indels 0; Gaps 0;

QY 51 GGCTGCGGCTGCGCGCGCGGCCCAATACGAACGCTACAGCTTCGCGAGCTTCCC 110
|||
Db 2849 GGCGCGGTGCTCGGATCGGCTCGNCACCTGGACAGCGCCGCGGTGGC 2908
|||
QY 111 ACGGACGAGCTGATCCGCTCGAGTCGGCTTACCGGCACGGCTGGACAAGTACAGCG 170
|||
Db 2909 GCTGGCGCGGAGTACTCTCGGGCTCGCGGACTGCGCGCATCGGCTCGCGGTGAC 2968
|||
QY 171 CGAGCACTGGCGCGAGCGCTGGCTACCTGGAGATCAGCTCGGCTGCACCGTTGCT 230
|||
Db 2969 CGGCGCCGACACCGACCCCGTCTGGCACCTCTTACCGTGGCAACGAGCGCGCGACA 3028
|||
QY 231 GCGCGACAGCGAGGCTTCTGCCACCGCAACTGCAAGCGCGCGCGCGAGCGCGCGC 290
|||
Db 3029 GCTGGCAGCCACCTCGACGCGCGCGGATCGACACCTCTACGCACTACCGGTACCGGT 3088
|||
QY 291 CGCGGCTTGGCGAGTATCCGAGTTCGGCTCTTTCGGGGGCTGCTGGCGCGCGCA 350
|||
Db 3089 GCACCTCTCGCGCGCTACGCGGCGAGGCAACCGCGGAAGGCTCGCTCCGCGGCGCA 3148
|||
QY 351 CTGCGCTCAAGCGGTGCAAGCAGGCGCTGCGAGCTTCGCGCAGTCCGAGCCGCGCGA 410
|||
Db 3149 GAGCTTCGCGCGCGAGGTCTCAGCTGCGGATCGGCGCGGACCTGGAGCGCGCGCAGGC 3208
|||
QY 411 GTGTGCGCGAGCTTCAGCGCGCGCA 437
|||
Db 3209 GCTGCGGTGATCGACGCGGTGCGCA 3235
|||

Search completed: January 30, 2004, 18:32:24
Job time : 69.8238 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2004, 07:55:54 ; Search time 2524.98 Seconds
(without alignments)
11579.600 Million cell updates/sec

Title: US-09-729-674-1_COPY_63_1265
Perfect score: 1203
Sequence: 1 atggagcggggcgccgggg.....tggaaactggagagaccagc 1203

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_htc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_htc:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: em_gss_hum:*
- 18: em_gss_inv:*
- 19: em_gss_pln:*
- 20: em_gss_vrt:*
- 21: em_gss_fun:*
- 22: em_gss_mam:*
- 23: em_gss_mus:*
- 24: em_gss_pro:*
- 25: em_gss_rod:*
- 26: em_gss_phg:*
- 27: em_gss_vrl:*
- 28: gb_gss1:*
- 29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1033.2	85.9	1201	9	AL545483
2	977.2	81.2	1201	9	AL514488
3	969.8	80.6	1201	9	AL556976
4	957.8	79.6	1201	13	BX425905

5	954	79.3	1148	13	BU902192
6	945.8	78.6	1075	9	AL546910
7	929.6	77.3	1201	9	AL541167
8	923.8	76.8	1656	11	AK047506
9	923	76.7	1201	9	AL517455
10	912.6	75.9	1652	11	AK017797
11	866.8	72.1	1201	9	AL543090
12	863.6	71.8	1201	9	AL554750
13	838	69.3	939	13	BX386576
14	834	69.3	1201	9	AL558186
15	822.2	68.3	1152	9	AL550966
16	805.8	67.0	1040	9	AL558966
17	801	66.6	846	13	BU838140
18	787.6	65.5	944	13	BU856662
19	784.4	65.2	1103	12	BM550012
20	771.4	64.1	774	9	AL598447
21	766.6	63.7	1083	12	BM920687
22	765.6	63.6	1201	13	BX387779
23	761.2	63.3	846	12	BI253401
24	760.4	63.2	899	13	BQ437797
25	756.4	62.9	1200	13	BU902214
26	752.2	62.5	904	13	BU931550
27	746.2	62.0	962	14	CD516948
28	740.4	61.5	878	13	BU183287
29	737	61.3	827	12	BI760650
30	735.4	61.1	1103	13	BQ233085
31	726	60.3	875	14	CD359485
32	720.6	59.9	1430	11	AK007624
33	718	59.7	1201	13	BX437872
34	717.2	59.6	912	13	BU175306
35	711.8	59.2	945	13	BQ677531
36	704	58.5	782	10	B8877856
37	702.6	58.4	1088	10	B8025893
38	702.4	58.4	837	13	BQ428356
39	700.8	58.3	1201	13	BX460044
40	700.2	58.2	707	10	BG747825
41	698	58.0	846	13	BU189067
42	697.8	58.0	713	10	B8746070
43	691.6	57.5	887	13	BQ718961
44	680.2	56.5	943	13	BU539494
45	674.8	56.1	943	10	BG323659

ALIGNMENTS

RESULT 1	AL545483	1201 bp	mRNA	linear	EST 31-MAY-2003
LOCUS	AL545483	Homo sapiens	PLACENTA COT 25-NORMALIZED	Homo sapiens	cdna
DEFINITION	clone CSODI015YK22 5-PRIME, mRNA sequence.				
ACCESSION	AL545483				
VERSION	AL545483.2	GI:31267318			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Li W.B., Gruber, C., Jessee, J. and Polayes, D.				
TITLE	Full-length cDNA libraries and normalization				
JOURNAL	Unpublished				
COMMENT	On Feb 15, 2001 this sequence version replaced gi:12877964. Contact: Genoscope Genoscope - Centre National de Sequencage Bp 191 91006 EVRY cedex - France Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 2233.f For more information about this cluster, see http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CSODI015Bf11QPI&cluster=2233.f. Contact : Feng Liang Email : fliang@lifetech.com URL :				

1	QY	ATGAGCCCGGGCGCCCGGGGGCGCGCGCCTCTAGCGCTGCTGTGCGTGCCTGCGCG	60
106	Db	ATGAGCCCGGGCGCCCGGGGGCGCGCGCCTCTAGCGCTGCTGTGCGTGCCTGCGCG	165
61	QY	CTGCGCGCGGGCGCGCCCAATACGAACGGTACAGCTTCCGAGCTTCCACAGGAGCAG	120
166	Db	CTGC - CGCGCGCGC - CCCAATACGAACGGTACAGCTTCCGAGCTTCCACAGGAGCAG	223
121	QY	CTGATGCGGCTCGAGTCGGCCTACCGGACGCGCTGGACAAGTACAGCGGCGAGCACTGG	180
224	Db	CTGATGCGGCTCGAGTCGGCCTACCGGACGCGCTGGACAAGTACAGCGGCGAGCACTGG	283
181	QY	GCCGAGAGCSTKGGCTACTCTGGAGATCAGCCTCGCGCTGCACGGTTGCTGCGGACAGC	240
284	Db	GCCGAGAGCGTGGGCTACTCTGGAGATCAGCCTCGCGCTGCACGGTTGCTGCGGACAGC	343
241	QY	GAGGCTTCTGCCACCCGAACCTGAGGCGCGCGCGCGACCCGAGCCCGCGCGCGGCTC	300
344	Db	-AGGCTTCTGCCACCCGAACCTGAGGCGCGCGCGCGACCCGAGCCCGCGCGCGGCTC	402
301	QY	GCCAGCTATCCGAGCTGCGCCTCTTCGGGGGGCTGCTCGCGCGCGCGCACTTGCTCAAG	360
403	Db	GCCAGCTATCCGAGCTGCGCCTCTTCGGGGGGCTGCTCGCGCGCGCGCACTTGCTCAAG	462
361	QY	CGCTGCAAGCAGGSCCTGCGAGCCTTTCGGCAGTCCCGCAGTCCAGCCGAGCGCGAGTGTGGCG	420
463	Db	CGCTGCAAGCAGGSCCTGCGAGCCTTTCGGCAGTCCCGCAGTCCCGCAGCGCGAGTGTGGCG	522
421	QY	GACTTCCAGCGCGCGAGCCCTACAAGTTCCTCGAGTTCGCTTACTCAAGGCAAAATAAT	480
523	Db	GACTTCCAGCGCGCGAGCCCTACAAGTTCCTCGAGTTCGCTTACTCAAGGCAAAATAAT	582
481	QY	CTCCCCAAGGCATCGCGCTGCTCACACTTTCTACTGAAGCATCTGTATGACGAATG	540
583	Db	CTCCCCAAGGCATCGCGCTGCTCACACTTTCTACTGAAGCATCTGTATGACGAATG	642
541	QY	ATGAAGAGGAACATGGCATATTATAAGAGCGCTCGCTGTGCGGAGGACTACATTAAAGAC	600
643	Db	ATGAAGAGGAACATGGCATATTATAAGAGCGCTCGCTGTGCGGAGGACTACATTAAAGAC	702
601	QY	CTGAARACCAAGTCATATGAAGCCTGTTTCATCGAGCAGTGTGGGGCATACAACGGTGAAG	660
703	Db	CTGAARACCAAGTCATATGAAGCCTGTTTCATCGAGCAGTGTGGGGCATACAACGGTGAAG	762
661	QY	AACCTGGAGAACATCCATCACAGACATGGAGCTGGGCCCTTCCGAGCTTCTCAAAGCCTTT	720
763	Db	AACCTGGAGAACATCCATCACAGACATGGAGCTGGGCCCTTCCGAGCTTCTCAAAGCCTTT	822
721	QY	TACAGTGTCTCGAGCCTCGAGGTTTCAGGGAGATCAAGGACTTCAAGGATTTCTAC	780
823	Db	TACAGTGTCTCGAGCCTCGAGGTTTCAGGGAGATCAAGGACTTCAAGGATTTCTAC	882


```

QY 661 AACTGGAGAAATCCATCACAGACATGAGGTGCGCCCTTCCGACTTCTTCAAAGCCTTT 720
Db 694 AACTGGAGAAATCCATCACAGACATGAGGTGCGCCCTTCCGACTTCTTCAAAGCCTTT 753
QY 721 TACGAGTGTCTCGAGAGCTCGAGGGTTCCAGGAGATCAAGGACTTCAAGGATTTCTAC 780
Db 754 TACGAGTGTCTCGAGAGCTCGAGGGTTCCAGGAGATCAAGGACTTCAAGGATTTCTAC 813
QY 781 CTTTCCATAGCAGATCAATATGTAGAAAGTTCTGGAATGCAAAATACAGTGTGAAGAGAAC 840
Db 814 CTTTCCATAGCAGATCAATATGTAGAAAGTTCTGGAATGCAAAATACAGTGTGAAGAGAAC 873
QY 841 CTCACCCAGATTAGAGAGCTATCCGGTTGAGAAATTTGGGTACCATGTATCATAC 900
Db 874 CTCACCCAGATTAGAGAGCTATCCGGTTGAGAAATTTGGGTACCATGTATCATAC 933
QY 901 TTGCAGTTTGCTATTATTAAGTTGAAGACCTGGAAGATGAGCCCTCTGTCAGTCAGC 960
Db 934 TTGCAGTTTGCTATTATTAAGTTG-ACGACCTGAAGATGAGCCCTCTGTCAGTCAGC 992
QY 961 TATCTGCTCTTTGATCAGATGACAGGTGATGAGCAGAACCTGGTGTATTACCATAC 1020
Db 993 TATCTGCTCTTTGATGAGAAATTAAGGKCATGAG-AGAACCTGGTGTATTACATKWM 1051
QY 1021 CACAGGGACACTTGGGGCTCTCGGATGAGCACTTCCAGCCCGAGACCTGAGCAGTTCAG 1080
Db 1052 MARGG---AMATGGGGSYTYGGT-KAVCTTYGACCCARACTTAACARTYAMMTYTT 1107
QY 1081 TTCTTTA 1087
Db 1108 TTWTTW 1114

RESULT 5
LOCUS BU902192
DEFINITION BU902192 1148 bp mRNA linear EST 17-OCT-2002
5', mRNA sequence.
AGENCOURT 10127451 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6502891
BU902192
BU902192.1 GI:24084105
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabbs-x@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M14059 row: k column: 20
High quality sequence stop: 754.
Location/Qualifiers
1..1148
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6502891"
/tissue_type="leiomysarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/notes="Organ: uterus; Vector: pCMV-Sport6; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.1 kb."
FEATURES
SOURCE

```

```

BASE COUNT 253 a 363 c 315 g 216 t 1 others
ORIGIN
Query Match 79.3%; Score 954; DB 13; Length 1148;
Best Local Similarity 96.2%; Pred. No. 2.8e-216;
Matches 1008; Conservative 2; Mismatches 32; Indels 6; Gaps 3;
QY 5 AGCCGGGGCGCGGGGGCGCGGCTGCTAGCGTGTGCTGGCTCGGCGCTGC 64
Db 1 AGCCGGGGCGCGGGGGCGCGGCTGCTAGCGTGTGCTGGCTCGGCGCTGC 60
QY 65 GGGCGGGGGCGCGCCCAATACGAACGCTACAGTTCCGAGCTTCCACGGGACGAGCTGA 124
Db 61 GGGCGGGGGCGCGCCCAATACGAACGCTACAGTTCCGAGCTTCCACGGGACGAGCTGA 120
QY 125 TGCCTGCTCAGTTCGCTTACCGGCACGCGCTGCAAGTACAGCGGCGGAGCAGCTGGGCGG 184
Db 121 TGCCTGCTCAGTTCGCTTACCGGCACGCGCTGCAAGTACAGCGGCGGAGCAGCTGGGCGG 180
QY 185 AGAGSTKGGTACCTGAGATCAGCTTCGCTGCGCTTGCACCGCTTGTGCGCGACGAGG 244
Db 181 AGAGCGTGGGTACCTGAGATCAGCTTCGCTGCGCTTGCACCGCTTGTGCGCGACGAGG 240
QY 245 CTTTCTGCCACGCAACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 304
Db 241 CTTTCTGCCACGCAACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
QY 305 GCTATCCGAGTGGCGCTCTTTCGGGGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCG 364
Db 301 GCTATCCGAGTGGCGCTCTTTCGGGGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCG 360
QY 365 GCAAGCAGGGCTGCGCGCGCTTCCGCGAGTCCGCGCGCGCGCGCGCGCGCGCGCGCGAG 424
Db 361 GCAAGCAGGGCTGCGCGCGCTTCCGCGAGTCCGCGCGCGCGCGCGCGCGCGCGCGCGAG 420
QY 425 TCCAGCGCGCGAGCGCTTACAAAGTTCTCGAGTTCGCTTACTTCAAGGCAAAATATCTCC 484
Db 421 TCCAGCGCGCGAGCGCTTACAAAGTTCTCGAGTTCGCTTACTTCAAGGCAAAATATCTCC 480
QY 485 CCAAGGCCATGCGCGCTGCTCACACCTTTCTACTGAAGCATCTCTGATGACGAAATGATGA 544
Db 481 CCAAGGCCATGCGCGCTGCTCACACCTTTCTACTGAAGCATCTCTGATGACGAAATGATGA 540
QY 545 AGAGAAATCGCATATTATTAAGAGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 604
Db 541 AGAGAAATCGCATATTATTAAGAGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
QY 605 AAACCAAGTCATATGAAGCGCTTTCATCCGAGCAGTGGCGGCATCAACCGTGAGAACT 664
Db 601 AAACCAAGTCATATGAAGCGCTTTCATCCGAGCAGTGGCGGCATCAACCGTGAGAACT 660
QY 665 GGAGAACATCCATCACAGACATGAGCTGCGCTTCCGAGCTTCTTCAAAGCCTTTTACG 724
Db 661 GGAGAACATCCATCACAGACATGAGCTGCGCTTCCGAGCTTCTTCAAAGCCTTTTACG 720
QY 725 AGTGTCTCGAGCTGCGAGGGTTCCAGGAGATCAAGGACTTCAAGGATTTCTACCTTT 784
Db 721 AGTGTCTCGAGCTGCGAGGGTTCCAGGAGATCAAGGACTTCAAGGATTTCTACCTTT 780
QY 785 CCATAGCAGATCATATTGTAGAAGTTCTGGAATGCAAAATACAGTGTGAAGAGAACTCA 844
Db 781 CCATAGCAGATCATATTGTAGAAGTTCTGGAATGCAAAATACAGTGTGAAGAGAACTCA 840
QY 845 CCCAGTTATAGGAGCTATCCGCTTGAGAAATTTGGGTACCATGTATCATCTACTTCG 904
Db 841 CCCAGTTATAGAGGCTATCCGCTTGAGAAATTTGGGTACCATGTATCATCTACTTCG 900
QY 905 AGTTTGCTATTATTAAGTTGAACGACCTGAAGATGCGAG-CCCCCTGTGTCAGTCAGTAT 963
Db 901 AGTTTGCTATTATTAAGTTGAACGACCTGAAGATGCGAG-CCCCCTGTGTCAGTCAGTAT 960
QY 964 CTGCTCTTTCATCAGATGACAGGTCATGCGAG--AGAACCTGGTGTATTACAGTACC 1021

```

```

Db      961 CGGCTCTTGTATCAGAAAGACAGGCTCTGACGCAACAACATGGGGGTATTACCGCTCC 1020
Qy      1022 ACAGGGACAC---TTGGGGCTCTTCGGA 1046
Db      1021 CCCCGGGACCCCTTTGGGGCCCTTCGGA 1048

RESULT 6
AL546910 1075 bp mRNA linear EST 31-MAY-2003
LOCUS     AL546910 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0DI032YH24 5-PRIME, mRNA sequence.
ACCESSION AL546910
VERSION   AL546910.2 GI:31268743
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1075)
AUTHORS   Li W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE     Full-length cDNA libraries and normalization
JOURNAL   Unpublished
COMMENT   On Feb 15, 2001 this sequence version replaced gi:12880487.
           Contact: Genoscope
           Genoscope - Centre National de Sequencage
           BP 191 91006 EVRY cedex - France
           Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
           Library was constructed by Life Technologies, a division of
           Invitrogen. This sequence belongs to sequence cluster 2233.f For
           more information about this cluster, see
           http://www.genoscope.cns.fr/
           cgi-bin/cluster.cgi?seq=CS0DI032DD12QPI&cluster=2233.f. Contact :
           Feng liang Email : fliang@lifetech.com URL :
           http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
           Faraday Avenue Genoscope sequence ID : CS0DI032DD12QPI.
FEATURES
     source
     1..1075
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone="CS0DI032YH24"
     /tissue_type="PLACENTA COT 25-NORMALIZED"
     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
     /note="First strand cDNA was primed with a NotI-oligo(dT)
     primer. Five prime end enriched, double-strand cDNA was
     digested with Not I and cloned into the Not I and EcoR V
     sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 240 a 317 c 296 g 212 t 10 others
ORIGIN

Query Match 78.6%; Score 945.8; DB 9; Length 1075;
Best Local Similarity 98.8%; Pred. No. 2.4e-214;
Matches 966; Conservative 6; Mismatches 4; Indels 2; Gaps 2;
Qy      2 TGGAGCGGGGCGCGGGGGCGCGCGCGCTGTAGCGCTGTGTGCGCTGTGCGCTGTGCGCG 61
Db      62 TGGAGCGGGGCGCGGGGGCGCGCGCGCTGTAGCGCTGTGTGCGCTGTGCGCTGTGCGCG 121
Qy      62 TGGCGCGCGGGGCGCGCGCGCTGTAGCGCTGTGTGCGCTGTGCGCTGTGCGCGCGAGC 121
Db      122 TGGCGCGCGGGCGCGCGCGCTGTAGCGCTGTGTGCGCTGTGCGCTGTGCGCGCGAGC 181
Qy      122 TGATCGCGCTGTAGCTGTGCGCTGTGCGCGCGCTGTGCGCGCGCTGTGCGCGCGAGCTGGG 181
Db      182 TGATCGCGCTGTAGCTGTGCGCTGTGCGCGCGCTGTGCGCGCGCTGTGCGCGCGAGCTGGG 241
Qy      182 CCGAGAGCTKGGCTACTGTGAGATCAGCTGTGCGCTGTGCGCGCGCTGTGCGCGAGCG 241
Db      242 CCGAGAGCTGTGGCTACTGTGAGATCAGCTGTGCGCTGTGCGCGCGCTGTGCGCGAGCG 301
Qy      242 AGGCTTTCTGCCACCGCAACTGTGAGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCTCG 301

```

```

Db      302 AGGCTTTGTCCACCGCAACTGTGACGCGCGCGCGCGAGCCCGCGCGCGCGCTCG 361
Qy      302 CAGCTATCCCGAGCTTGGCGCTTTCGGGGGCGCTGTGCGCGCGCGCGAGCTGTGCTCAAGC 361
Db      362 CAGCTATCCCGAGCTTGGCGCTTTCGGGGGCGCTGTGCGCGCGCGCGAGCTGTGCTCAAGC 421
Qy      362 GTTCAAGCAGGCGCTTGGCGCGCTTTCGGCGAGTCCCGAGCCCGAGCGCGCGAGGTGTGGCG 421
Db      422 GCTCAAGCAGGCGCTTGGCGCGCTTTCGGCGAGTCCCGAGCCCGAGCGCGCGAGGTGTGGCG 481
Qy      422 ACTTCAGCGCGCGCGAGCGCTTCAAGTTCCTCGAGTTCGCTTACTTCAAGGCAAAATAATC 481
Db      482 ACTTCAGCGCGCGCGAGCGCTTCAAGTTCCTCGAGTTCGCTTACTTCAAGGCAAAATAATC 541
Qy      482 TCCCAAGCCATCGCGCGCTGTCTACACCTTCTTACTGAAGCATCTGTATGACGAAATGA 541
Db      542 TCCCAAGCCATCGCGCGCTGTCTACACCTTCTTACTGAAGCATCTGTATGATGAATGA 601
Qy      542 TGAAGAGAACATGGCATATTATAAGAGCCTCCCTGGTGCAGGAGTACATTAAAGACC 601
Db      602 TGAAGAGAACATGGCATATTATAAGAGCCTCCCTGGTGCAGGAGTACATTAAAGACC 661
Qy      602 TGAAGAACCAAGTCATATGAAGCCCTTTCATCCGAGCGAGTGGCGGCATACACGGTGAGA 661
Db      662 TGAAGAACCAAGTCATATGAAGCCCTTTCATCCGAGCGAGTGGCGGCATACACGGTGAGA 721
Qy      662 ACTGGAGAACATCCATCAGACATGAGAGTGGCGCTTCCCGACTTCTTCAAGGCTTTT 721
Db      722 ACTGGAGAACATCCATCAGACATGAGAGTGGCGCTTCCCGACTTCTTCAAGGCTTTT 781
Qy      722 AGAGTGTCTCGACGCTCGGAGGTTCCAGGAGATCAAGGACTTCAAGGATTTCTTACC 781
Db      782 AGAGTGTCTCGACGCTCGGAGGTTCCAGGAGATCAAGGACTTCAAGGATTTCTTACC 841
Qy      782 TTTCATAGCAGATCATTTATGTAGAGTTCTGGAATGCAAAATACAGTGTGAAGAGAAC 841
Db      842 TTTCATAGCAGATCATTTATGTAGAGTTCTGGAATGCAAAATACAGTGTGAAGAGAAC 901
Qy      842 TCACCCCGAGTTATAGAGGCTATCCGGTTGAGAAATTTGTGGTACCATGTATCATTTACT 901
Db      902 TCACCCCGAGTTATAGAGGCTATCCGGTTGAGAAATTTGTGGTACCATGTATCATTTACT 961
Qy      902 TGCAGTTTGCCTATTATTAAGTTGAACGACCTTGAAGATGCAAGTCCCGCGAGCCCGCTGTGCGTCACT 961
Db      962 TGCAGTTTGCCTATTATTAAGTTGAACGACCTTGAAGATGCAAGTCCCGCGAGTCCCGTCACT 1019
Qy      962 ATCTGCTCTTTGATCAGA 979
Db      1020 ATCTGCTCTTTGATMRAA 1037

```

```

RESULT 7
AL541167 1201 bp mRNA linear EST 12-MAY-2003
LOCUS     AL541167 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0E005YF12
DEFINITION 5-PRIME, mRNA sequence.
ACCESSION AL541167
VERSION   AL541167.2 GI:30545079
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS   Li W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE     Full-length cDNA libraries and normalization
JOURNAL   Unpublished
COMMENT   On Feb 15, 2001 this sequence version replaced gi:12871971.
           Contact: Genoscope
           Genoscope - Centre National de Sequencage
           BP 191 91006 EVRY cedex - France
           Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
           Library was constructed by Life Technologies, a division of

```

Invitrogen. This sequence belongs to sequence cluster 2233.f For more information about this cluster, see <http://www.genoscope.cns.fr/csg-bin/cluster.cgi?seq=CS0DE05DC06P1&cluster=2233.f>. Contact : Feng Liang Email : liang@lifetech.com URL : <http://fulllength.invitrogen.com> / Invitrogen Corporation 1600 Faraday Avenue, Invitrogen sequence ID : CS0DE005DC060P1

FEATURES

```

location/qualifiers
1. 1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DB005VF12"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: PCMVSPORT6; 1st strand cDNA with a NotI-oligo(dT) primer. Five prime double-strand cDNA was digested with NotI and EcoRV. The Not I and EcoRV sites of the PCMVSPORT6 Library was not normalized."
258 a 335 c 313 g 237 t 58 others
BASE COUNT
ORIGIN

```

Query March	77.3%;	Score 929.6;	DB 9;	Length 1201;
Best Local Similarity	98.2%;	Pred. No. 1.8e-210;		
Matches 970;	Conservative 10;	Mismatches 4;	Indels 4;	Gaps 4;
Qy	1	ATGAGAGCCGGGCGCCGGGGGGCGCGCGCTGCTAGCGCTGTGTCGCTGGCTGGCTGGCGG	60	
Db	64	ATGAGAGCCGGGCGCCGGGGGGCGCGCGCTGCTAGCGCTGTGTCGCTGGCTGGCTGGCGG	123	
Qy	61	CTGCGCGCCGGGCGCGGCCCAATACGAACGCTACAGCTTCGCGACGCTCCCCACGGGACGAG	120	
Db	124	CTGCGCGCGCGGCGC - CCCAATACGAACGCTACAGCTTCGCGAGCTCCACCGGACGAG	182	
Qy	121	CTGATCGCGCTCGAGTCGGCTACCGCCACGCGCTGGAACAAGTACAGCGCGAGCACTGG	180	
Db	183	CTGATCGCGCTCGAGTCGGCTACCGCCACGCGCTGGACAAGTACAGCGCGAGCACTGG	242	
Qy	181	GCCGAGAGCSTKGCTACCTCGGAGATCAGAGCTCGGCTGCAACGCTGTGTCGCGCACAGC	240	
Db	243	GCCGAGAGCTGGGCTACCTGGAGATCAGAGCTCGGCTGCAACGCTGTGTCGCGCACAGC	302	
Qy	241	GAGGCTTCTGCCACCGCAACTGAGCGCGCGCGGAGCCGAGCCGCGCGCTCCGCGGCTTC	300	
Db	303	-AGGCTTCTGCCACCGCAACTGAGCGCGCGCGGAGCCGAGCCGCGCGCTCCGCGGCTTC	361	
Qy	301	GCCAGCTATCCCGAGCTGCGCCTCTTCGGGGGGCTGCTGGCGCGCGGCACTGCCTCAAG	360	
Db	362	GCCAGCTATCCCGAGCTGCGCCTCTTCGGGGGGCTGCTGGCGCGCGGCACTGCCTCAAG	421	
Qy	361	CGTGTCAAGCAGGGCTGCGAGCCTTCGCGCAGTCCCGACGCCAGCGCGAGTGTGTGGCG	420	
Db	422	CGTGTCAAGCAGGGCTGCGAGCCTTCGCGCAGTCCCGACGCCAGCGCGAGTGTGTGGCG	481	
Qy	421	GACTTCCAGCGCGCGAGCCCTACAGTTCTGTGAGTTCTGCTTACTTCAAGGCAATAAT	480	
Db	482	GACTTCCAGCGCGCGAGCCCTACAGTTCTGTGAGTTCTGCTTACTTCAAGGCAATAAT	541	
Qy	481	CTCCCCAAGCCATCGCCCGCTGCTCACACCTTCTACTGAAGCATCTCTATCAGCAATG	540	
Db	542	CTCCCCAAGCCATCGCCCGCTGCTCACACCTTCTACTGAAGCATCTCTATGATGAATG	601	
Qy	541	ATGAAGAGGAACAATGCGATATTATTAAGAGCGCTGCTGGTGGCGAGACTACATTAAAGAC	600	
Db	602	ATGAAGAGGAACAATGCGATATTATTAAGAGCGCTGCTGGTGGCGAGACTACATTAAAGAC	661	
Qy	601	CTGGAAACCAAGTCATATGAAGCCTGTTCATCCGAGCAGTGCGGGCATACACGGTCAG	660	
Db	662	CTGGAAACCAAGTCATATGAAGCCTGTTCATCCGAGCAGTGCGGGCATACACGGTCAG	721	
Qy	661	AACTGTGAGAACATCCCATCAAGACATGGAGCTGGCCCTTCGCGACTTCTTCAAGGCCCTT	720	

Db	722	AATGGAGAACATCCATCACAGACATGGAGCTGGCCCTCCCGACTTCCTTCAAAGCCTTT	781
Qy	721	TACGAGTGTCTCGAGCGCTCGCAGGGGTTCAGGAGATCAAGGACTTCAAGGATTTCTAC	780
Db	782	TACGAGTGTCTCGAGCGCTCGCAGGGGTTCAGGAGATCAAGGACTTCAAGGATTTCTAC	841
Qy	781	CTTTCATACAGATCATATGTAGTAAGTCTCTGGAATGCAAAATACAGTGTGAAGAAC	840
Db	842	CTTTCATACAGATCATATGTAGTAAGTCTCTGGAATGCAAAATACAGTGTGAAGAAC	901
Qy	841	CTCACCCAGTATAGGAGGCTATCCGGTTGAGAAATTTGTGGCTACCATGTATCATTC	900
Db	902	CTCACCCAGTATAGGAGGCTATCCGGTTGAGAAATTTGTGGCTACCATGTATCATTC	961
Qy	901	TTGAGTTTCCCTATTATTAAGTTGAACGCTGGAAGATGCAGCCCCCTGTGCAGTCAGC	960
Db	962	TTGAGTTTCCCTATTATTAAGTTGAACGCTGGAAGATGCAGCCCCCTGTGCAGTCAGC	1019
Qy	961	TATCTGCTCTTTGATCAGATGACRAGG	988
Db	1020	TATCTGCTCTTTGATCAGATGACRAGG	1047
RESULT 8			
AK047506			
LOCUS		1656 bp mRNA linear HTC 05-DEC-2002	
DEFINITION		Mus musculus 10 days neonate cerebellum cDNA, RIKEN full-length enriched library, clone:B930079M13 product:cartilage associated protein, full insert sequence.	
ACCESSION		AK047506	
VERSION		AK047506.1 GI:26338809	
KEYWORDS		HTC; CAP trapper.	
SOURCE		Mus musculus (house mouse)	
ORGANISM		Mus musculus	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;	
AUTHORS		1	
TITLE		Carninci, P. and Hayashizaki, Y.	
JOURNAL		High-efficiency full-length cDNA cloning	
MEDLINE		Meth. Enzymol. 303, 19-44 (1999)	
PUBMED		99279253	
REFERENCE		10349636	
AUTHORS		2	
TITLE		Shibata, K., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.	
JOURNAL		Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes	
MEDLINE		Genome Res. 10 (10), 1617-1630 (2000)	
PUBMED		20499374	
REFERENCE		11042159	
AUTHORS		3	
TITLE		Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kasinagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, K., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.	
JOURNAL		RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer	
MEDLINE		Genome Res. 10 (11), 1757-1771 (2000)	
PUBMED		20530913	
REFERENCE		11076861	
AUTHORS		4	
TITLE		Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schrim, L.M., Stauber, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,	

Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
Tayo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
and Hayashizaki,Y.

TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409 (6821), 695-690 (2001)

MEDLINE 21085660

PUBMED 11217851

REFERENCE

AUTHORS

5 The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

REFERENCE

AUTHORS

6 (bases 1 to 1656)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,K., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.

Direct Submission

TITLE Submitted (16-JUL-2001) Yoshinide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

Please visit our web site for further details.

URL:http://genome.gsc.riken.go.jp/

URL:http://fantom.gsc.riken.go.jp/

FEATURES

source

1..1656
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:B930079M13"
/db_xref="taxon:10090"
/clone="B930079M13"
/tissue_type="cerebellum"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stages="10 days neonate"
70..1272

CDS

/note="unnamed protein product; cartilage associated
protein (MGD|WGI:1891221, GI|NW_019922, evidence: BLASTN,
99%, match=1655)
putative"

/codon_start=1

/protein_id="BAC33076.1"

/db_xref="GI:26338810"

/translation="WGPRSPPTALLVLCVCAPTPOGQYERYSFSPFRDELMPLE
SAYRALDQISGEHWASVGLVSLRLHLLRDSFAFCHRNCSAATPAPAPGASH
AELRFGVLRRACLRKQGLPAFQSPQSRVLAADFOOREFYFLQPAYFKANDL
PKAAAHYLLKHPGDDMMKRNMEYKSLFGAEDHKILETKSYSLFVRAVYNG
ENWRTSISDMELALPDFLKAIFYELAAACEGSRREIKDFKDFYLSIADHYVEVLECKRC

EEETLPVIGGYPVEKFAVTWYHYLOFAYYKLNDAKNAAPCAVSYLLPDDSDRVMQONL
VYQYHRDKWGLSDSEHFQRPPEAVQFNVNTTLQKELYDFAQEHLMDDDEGEVVEYVDD
LLETSSA"

BASE COUNT 341 a 498 c 448 g 369 t

ORIGIN

Query Match 76.8%; Score 923.8; DB 11; Length 1656;
Best Local Similarity 86.0%; Pred. No. 5.2e-209;
Matches 1021; Conservative 2; Mismatches 164; Indels 0; Gaps 0;

Qy 13 CGCGGGGGCCGCGCGCTGTAGCGCTGTGTGCGTGTGCGCTGCGCGCTGCGCGCGG 72
Db 79 CGCAGCCCTACGCGCCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 138
Qy 73 CGCGCCCAATACCAAGCTACAGCTTCCGAGCTTCCACGGGACGAGCTGATCGCGCTC 132
Db 139 CGCGGGCAGTACGAGCGCTACAGCTTCCGAGCTTCCGCGGGGACGAGCTGATCGCGCTC 198
Qy 133 GAGTCGGCTTACCGGACGCGCTTGGACAAGTACAGCGGCGAGCACTGGCGCCGAGAGCSTK 192
Db 199 GAGTCGGCTTACCGGACGCGCTTGGACAAGTACAGCGGCGAGCACTGGCGCCGAGAGCSTG 258
Qy 193 GGCTACCTGGAGATCAGCTTGGCGCTGACCGCTTCTGCTGGCGACAGCGAGCCTTCTCG 252
Db 259 GGCTACCTGGAGTGTAGCTTGGCGCTGACCGCTTCTGCTGGCGACAGCGAGCCTTCTCG 318
Qy 253 CACGCAACTGACGCGCGCGCGCGCGAGCGCGCGCGCGCGCTGCCAGCTATCCC 312
Db 319 CACGCAACTGACGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 378
Qy 313 GAGTCGGCTTCTGGGCGCGCTGCTGCGCGCGCGCTGCTGCTGAGCGCTGCGAGCGT 372
Db 379 GAGTACGCTTCTGGCGCGCTGCTGCGCGCGCGCTGCTGCTGAGCGCTGCGAGCGT 438
Qy 373 GGCTGCCAGCTTCCGCGAGTCCACAGCGCGAGCGTGTCTGCGCGACTTCCAGCGC 432
Db 439 GGCTGCCAGCTTCCGCGAGTCCACAGCGCGAGCGCTGCTGCTGCGCGACTTCCAGCGC 498
Qy 433 CGCGAGCGCTTACAGTCTTCTGCGAGTCTGCTTCTGCTGAGCGAAATATCTCCCAAGCC 492
Db 499 CGCGAGCGCTTACAGTCTTCTGCGAGTCTGCTTCTGAGCGCAATGACCTCCGAGGCG 558
Qy 493 ATCGCGCTGCTCACACCTTTTCTATGAAGCATCTCTGATGAGCAATGATGAAGAGAAC 552
Db 559 ATCGTGGCGTCAACACCTATCTCTGAGCATCCAGATGACGAGTGTGAGAGAAC 618
Qy 553 ATGCGATATTAAGAGCGCTGCTGCTGCGGAGGACTACATTAAAGACCTGGAACCAAG 612
Db 619 ATGAGTATTTAAGAGCTTGCCTGGAGCGGAGGACCAATTAAAGACTTGGAAACCAAG 678
Qy 613 TCATATGAAGCGCTGTTCATCCGAGCGAGTGGCGGATACAACGGTGAGAACTGGAGAAC 672
Db 679 TCGTACGAGCGCTGTTCCTGCGTGGTGGCGGCTACACGCGGAGACTTGGAGAAC 738
Qy 673 TCATCACAGACATGGAGCTGCGCTTCCGACTTCTTCAAAGCCTTTTACAGTGTCTC 732
Db 739 TCATTTCCGACATGGAGCTGCGCTTCCGACTTCTTCAAAGCCTTTTACAGTGTCTC 798
Qy 733 GCAGCTGCGGCGTTCACAGGAGATCAAGGACTTCAAGGATTTCTACCTTTCCATAGCA 792
Db 799 GCTGCTGCGAGGGGTGCGGAGATCAAGGACTTCAAGGACTTCTACTGTTCCATAGCA 858
Qy 793 GATCATTTGTAGAAGTCTTGGAAATGCAAAATACAGTGTGAAGAGAACCTCACCCAGTT 852
Db 859 GATCATTTGTGAAGTCTTGGAGTGAAGTTCGTTGTGAGGAGACCTCACCCAGTCTC 918
Qy 853 ATAGAGGGCTATCCGCTTGAGAAATTTCTGGCTACCATGTATCATTTACAGTTTGGC 912
Db 919 ATAGAGGGCTATCCCGTGAGAAATTTTGGCGACCATGTACCATTTATTACAGTTTGGC 978
Qy 913 TATTATAAGTTGAACGAGCTGAAGAAATGAGGAAATGAGGAAATGAGGAAATGAGGAAAT 972
Db 979 TATTACAAGTTGAATGATCTGAAGATGAGGAAATGAGGAAATGAGGAAATGAGGAAAT 1038

QY 973 GATCAGATGACAGAGTCTATGACAGAGAACTCTGGTGTATATACAGTACCAAGGACACT 1032
 Db 1039 GACCAAGATGACAGAGTCTATGACAGAGAACTCTGGTGTATATACAGTACCAAGGACACT 1098
 QY 1033 TGGGGCCCTCTCGAATGAGCACTTCCAGCCAGAGCTGAAGAGTTCAGTTCTTTAATGTG 1092
 Db 1099 TGGGGCCCTCTCGAATGAGCACTTCCAGCCAGAGCTGAAGAGTTCAGTTCTTTAATGTG 1158
 QY 1093 ACCACACTCCAGAGAGCTGTATGACATTTCTTGAAGAAATATATATGATGATGATGAG 1152
 Db 1159 AGCAGCTCCAGAGAGCTGTATGACATTTCTTGAAGAAATATATATGATGATGATGAG 1218
 QY 1153 GGAGAAGTGTGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1199
 Db 1219 GGAGAAGTGTGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1265

 RESULT 9.
 ALS17455
 LOCUS
 DEFINITION ALS17455 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
 CS0DA003YF11 5-PRIME, mRNA sequence.
 ACCESSION ALS17455
 VERSION ALS17455.2 GI:30534775
 KEYWORDS EST.
 ORGANISM Homo sapiens (human)
 SOURCE
 REFERENCE
 AUTHORS Li W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 COMMENT On Feb 13, 2001 this sequence version replaced gi:12780948.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 2233.f For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DA003YF11&cluster=2233.f. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Paraday Avenue, Genoscope sequence ID : CS0DA003YF11.
 Location/Qualifiers
 1. 1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DA003YF11"
 /tissue_type="NEUROBLASTOMA"
 /clone_lib="Homo sapiens NEUROBLASTOMA"
 /note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
 with a NotI-oligo(dT) primer. Five prime end enriched,
 double-strand cDNA was digested with Not I and cloned into
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.
 Library was not normalized."
 BASE COUNT 270 a 327 c 306 g 240 t 58 others
 ORIGIN

Query Match 76.7%; Score 923; DB 9; Length 1201;

Best Local Similarity 97.3%; Pred. No. 6.8e-209;

Matches 956; Conservative 13; Mismatches 8; Indels 6; Gaps 3;

QY 6 GCGGGGCGCGGGGGCGCGGGCGCTGCTAGCGTGTGTGCGTGGCGCTGCGCGCTGCG 65

Db 65 GCGGGGCGCGGGGGCGCGGGCGCTGCTAGCGTGTGTGCGTGGCGCTGCGCGCTGCG 124

QY 65 GCGGGGCGCGGGGGCGCGGGCGCTGCTAGCGTGTGTGCGTGGCGCTGCGCGCTGCG 125

Db 125 CCGGGCGG---CCATACGAAACGCTACAGCTTCCGACAGCTTCCACGGACAGAGCTGAT 181
 QY 126 GCGGCTCGAGTCGCGCTTACCGGACCGGCTGGACAAGATACAGCGGAGACACTTGGGCGGA 185
 Db 182 GCGGCTCGAGTCGCGCTTACCGGACCGGCTGGACAAGATACAGCGGAGACACTTGGGCGGA 241
 QY 186 GAGCTGCGCTTACCTGGAGATCAGCTTGGGCTGGACCGCTTCTGCGGACAGCGAGGC 245
 Db 242 GAGCTGCGCTTACCTGGAGATCAGCTTGGGCTGGACCGCTTCTGCGGACAGCGAGGC 299
 QY 246 CTTCTGCCACCGCAACTGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 305
 Db 300 CTTCTGCCACCGCAACTGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 359
 QY 306 CTATCCGAGCTGCGCTTCTGCGGCGCGCTTCTGCGGCGCGCGCGCGCGCGCGCGCGCGCTG 365
 Db 360 CWTCCCGAGCTGCGCTTCTTCCGGGCGCTTCTGCGGCGCGCGCGCGCGCGCGCGCGCTG 419
 QY 366 CAAGCAGGCGCTGCGAGCGCTTCCGCGAGTCCGACCGCGCGCGCGCGCGCGCGCGCGCGCTT 425
 Db 420 CAAGCAGGCGCTGCGAGCGCTTCCGCGAGTCCGACCGCGCGCGCGCGCGCGCGCGCGCGCTT 479
 QY 426 CCAGCG 485
 Db 480 CCAGCG 539
 QY 486 CAAAGCCATCGCGCGCTGCTCAGACCTTCTTACTGAAGCATCTCTGATGACGAATGATGA 545
 Db 540 CAAAGCCATCGCGCGCTGCTCAGACCTTCTTACTGAAGCATCTCTGATGACGAATGATGA 599
 QY 546 GAGGAACATGCGCATATATATAGAGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 605
 Db 600 GAGGAACATGCGCATATATATAGAGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 659
 QY 606 AACCAAGTCATATGAAAGCGCTTCTTCTCCGAGCAGTCCGGCGCATACAAACGCTGAGAACTG 665
 Db 660 AACCAAGTCATATGAAAGCGCTTCTTCTCCGAGCAGTCCGGCGCATACAAACGCTGAGAACTG 719
 QY 666 GAGACATCCATCAGACATGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 725
 Db 720 GAGACATCCATCAGACATGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 779
 QY 726 GTGTCTCGAGCTCGGAGGCTTCCAGGAGAGATCAAGGACTTCAAGGATTTCTACCTTTC 785
 Db 780 GTGTCTCGAGCTCGGAGGCTTCCAGGAGAGATCAAGGACTTCAAGGATTTCTACCTTTC 839
 QY 786 CATAGCAGATCATTTATGATGAGTCTTGGATGCAAAATACAGTGTGAAGAGAACCTCAC 845
 Db 840 CATAGCAGATCATTTATGATGAGTCTTGGATGCAAAATACAGTGTGAAGAGAACCTCAC 899
 QY 846 CCCAGTTATAGAGGCTATCCGTTGAGAAATTTG-TGGTACCATGTATCATTTCTTTC 904
 Db 900 CCCAGTTATAGAGGCTATCCGTTGAGAAATTTGTTGGTACCATGTATCATTTCTTTC 959
 QY 905 AGTTTGCCTATATAGTGTGACACCTGAGATGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCTC 964
 Db 960 AGTTTGCCTATATAGTGTGACACCTGAGATGCGAGCGCGCGCGCGCGCGCGCGCGCGCTC 1019
 QY 965 TGCTCTTTGATCAGAAATGACAAG 987
 Db 1020 TGCTCTTTKATCAGAAKRAAAGG 1042

RESULT 10

AK017797

LOCUS

DEFINITION

AK017797

Accession

Version

Keywords

Source

AK017797 1652 bp mRNA linear HTC 05-DEC-2002
 Mus musculus 8 days embryo whole body cDNA, RIKEN full-length
 enriched library, clone:570529N23 product:cartilage associated
 protein, full insert sequence.

AK017797

AK017797.1 GI:12857226

HTC; CAP trapper.

Mus musculus (house mouse)


```
Db 487 CCAGCGCGCGGAGCCCTCAAGTTCCTCGAGTTCCTGCTTACTTCAAGGAAATAATCTCC 546
QY 486 CAAAGCCATCGCGCTGCTCACACCTTTCTACTGAAGCATCTGATGACGAATGATGAA 545
Db 547 CAAAGCCATCGCGCTGCTCACACCTTTCTACTGAAGCATCTGATGATGATGATGAA 606
QY 546 GAGGAACATGGCATATTAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 605
Db 607 GAGGAACATGGCATATTAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 666
QY 606 AACCAGTCTATGAAGCTGTTTCATCGAGCAGTGGCGGCGATACAAACGCTGAGAACTG 665
Db 667 AACCAGTCTATGAAGCTGTTTCATCGAGCAGTGGCGGCGATACAAACGCTGAGAACTG 726
QY 666 GAGAACATCCATCAGACATGAGTGGCTTCCGACCTTCCGACCTTCTTCAAGGCTTTTACGA 725
Db 727 GAGAACATCCATCAGACATGAGTGGCTTCCGACCTTCCGACCTTCTTCAAGGCTTTTACGA 786
QY 726 GTGCTCGAGCTCGAGGCTTCCAGGAGATCAAGGACTCAAGGACTTCAAGGACTTCTACCTTTC 785
Db 787 GTGCTCGAGCTCGAGGCTTCCAGGAGATCAAGGACTTCAAGGACTTCTACCTTTC 846
QY 786 CATACAGATCAATTAAGTGAAGTCTCGAATGCAAAATACAGTGTGAAGAACTTCAC 845
Db 847 CATACAGATCAATTAAGTGAAGTCTCGAATGCAAAATACAGTGTGAAGAACTTCAC 906
QY 846 CCCAGTATAGGAGCTATCCGCTTGAGAAATTTCTGCTCCTACATGATCAATTAAGTGA 905
Db 907 CCCAGTATAGG-RCATTCGCTTGAAG-ATTTGCTGCTACCATGATCAATTAAGTGA 964
QY 906 GTTTCCTTATTAAGTGA 925
Db 965 GTTTC-CTATTATAAGTGA 983
```

```
RESULT 12
AL554750
LOCUS
DEFINITION
AL554750 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1086YJ04 5-PRIME, mRNA sequence.
AL554750
VERSION
AL554750.2 GI:31276560
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 15, 2001 this sequence version replaced gi:12895833.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 2233.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0D1086DE02QPI&cluster=2233.f. Contact :
Feng Liang Email : liang@lifetech.com URL :
http://fulllength.invitrogen.com/ invitroGen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0D1086DE02QPI.
Location/Qualifiers
1. 1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon.9606"
/clone="CS0D1086YJ04"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)"
FEATURES
source
```

```
primer. Five prime end enriched, double-strand cDNA was
digested with NotI and cloned into the NotI and EcoRV
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 272 a 329 c 250 t 41 others
ORIGIN
Query Match 71.8%; Score 863.6; DB 9; Length 1201;
Best Local Similarity 95.3%; Pred. No. 9.5e-195;
Matches 945; Conservative 18; Mismatches 18; Indels 11; Gaps 7;
QY 1 ATGAGAGCGGGGCGCGGGGGCGCGGGCGCTGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
Db 77 ATGAGAGCGGGGCGCGGGGGCGCGGGGGCGCGGGCGCTGCTAGCTGCTGCTGCTGCTGCTGCT 136
QY 61 CTGCGGCGCGGGGCGCGCGCCCAATACGAACGCTACAGC- TTCCGAGCTTCCCAAGGAGCA 119
Db 137 CTGC-CGCGGGGCGC-CCCAATACGAACGCTACAGGCTTCCGAGCTTCCCAAGGAGCA 194
QY 120 GCTGATGCGGCTCGAGTCTGCGCTTACCGCAACGCGTGGACAAAGTACAGCGGAGCACTG 179
Db 195 GCTGATGCGGCTCGAGTCTGCGCTTACCGCAACGCGTGGACAAAGTACAGCGGAGCACTG 254
QY 180 G--GCCGAGAGGCTGGCTACCTGAGATCAGCTCGGCTGACGGCTGACGGCTGCTGCGGAC 237
Db 255 GCGCGGAGAGGCTGGCTACCTGAGATCAGCTGCGGCTGACGGCTGCTGCGGCGAC 314
QY 238 AGCGAGGCTTCTGCCACCGCAACTGCAAGCGCGCGCGCGCGAGCC--GCCGCGC 295
Db 315 ACTAKTCTCTTGGCACCGCAACTTCACTCCGCGCGCAKCCCKARCCCTGTCCGCG 374
QY 296 GCTCGCCAGCTATCCGAGCTCGGCTTCTCGGGGGCTGCTGCGCGCGGCACTGCGC 355
Db 375 GCTCGCCAGCTATCCGAGCTCGGCTTCTCGGGGGCTGCTGCGCGCGGCACTGCGC 434
QY 356 TCAAGCGCTGCAAGCAGGCGCTTCCGAGCTTCCGAGCTTCCGAGCTTCCGAGCTTCCGAGCTG 415
Db 435 TCAAGCGCTGCAAGCAGGCGCTTCCGAGCTTCCGAGCTTCCGAGCTTCCGAGCTTCCGAGCTG 494
QY 416 TGGCGAGCTTCCAGCGCGCGGAGCCCTCAAGTTCCTGAGTTCGCTTACTTCAAGGAA 475
Db 495 TGGCGAGCTTCCAGCGCGCGGAGCCCTCAAGTTCCTGAGTTCGCTTACTTCAAGGAA 554
QY 476 ATAATCTCCCCAAAGCCATCGCGCTGCTCACACCTTCTACTGAAGCATCTGATGACG 535
Db 555 ATAATCTCCCCAAAGCCATCGCGCTGCTCACACCTTCTACTGAAGCATCTGATGATG 614
QY 536 AAATGATGAAGAGAACATGCGATATTATAAGAGCTTCCCTGGTGGCGGAGCACTACATTA 595
Db 615 AAATGATGAAGAGAACATGCGATATTATAAGAGCTTCCCTGGTGGCGGAGCACTACATTA 674
QY 596 AAGACTCTGAAACCAAGTCATATGAAGGCTTTCATCCGAGAGCTGCGGCGCATACACG 655
Db 675 AAGACTCTGAAACCAAGTCATATGAAGGCTTTCATCCGAGAGCTGCGGCGCATACACG 734
QY 656 GTGAGAACTGAGAACATCCATCAAGAGCTGAGCTTCCGCTTCCGAGCTTCTTCAAG 715
Db 735 GTGAGAACTGAGAACATCCATCAAGAGCTGAGCTTCCGCTTCCGAGCTTCTTCAAG 794
QY 716 CTTTTCAGAGTGTCTGCGACCTTCCAGGGTTCAGGGAGATCAAGAGCTTCAAGATT 775
Db 795 CTTTTCAGAGTGTCTGCGACCTTCCAGGGTTCAGGGAGATCAAGAGCTTCAAGATT 854
QY 776 TCTACCTTTCATAGCAGATCATTTATGAGAGTTCGGAATGCAAAATACAGTGTGAAG 835
Db 855 TCTACCTTTCATAGCAGATCATTTATGAGAGTTCGGAATGCAAAATACAGTGTGAAG 914
QY 836 AGAACCTCACCCAGTTATAGGAGCTATCCGGTTGAGAAATTTGGCTTACCATGTATC 895
Db 915 AGAACCTCACCCAGTTATAGGAGCTATCCGGTTGAGAAATTTGGCTTACCATGTATC 974
QY 896 ATTACTTGCAGTTTGCCTATTATAAGTTGAACGACTGAGAGATGCGAGCCCCCTGTGCG 955
Db 975 ATTACTTGCAGTTTGCCTATTATAAGTTGAACGACTGAGAGATGCGAGCCCCCTGTGCG 1031
```



```

QY 956 TCAGTATCTGCTCTTTGATCAGATGACAAG 987
DB 1032 TMA-STAFCTCTBTGATCAGATGACARG 1062

RESULT 13
LOCUS BX386576 339 bp mRNA linear EST 08-MAY-2003
DEFINITION BX386576 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
ACCESSION BX386576 Homo sapiens cDNA clone CS0DJ007YB14 5-PRIME, mRNA sequence.
VERSION BX386576.1 GI:30455453
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 939)
Li, W.B., Gruber, C., Jesses, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 2233.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS1AJ002ZF08QP1&cluster=2233.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS1AJ002ZF08QP1.

FEATURES
source
1..939
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ007YB14"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
10-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 201 a 291 c 271 g 158 t 18 others
ORIGIN
Query Match 69.7%; Score 838; DB 13; Length 939;
Best Local Similarity 97.7%; Pred. No. 1e-188;
Matches 840; Conservative 17; Mismatches 2; Indels 1; Gaps 1;

QY 1 ATGAGCGCGGGCGCGCGGGCGCGCGCGCTGCTAGCGCTGCTGCGCTGCGCTGCGG 60
DB 81 ATGAGCGCGGGCGCGCGGGCGCGCGCGCTGCTAGCGCTGCTGCGCTGCGCTGCGG 140
QY 61 CTGCGCGCGCGGGCGCGCGCGCGCGCGCTGCTAGCGCTGCTGCGCTGCGCTGCGG 120
DB 141 CTGCGCGCGCGGGCGCGCGCGCGCGCGCGCTGCTAGCGCTGCTGCGCTGCGCTGCGG 200
QY 121 CTGATCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGG 180
DB 201 CWTGTCGGCYCGAGTCGGCTTACCGGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGG 260
QY 181 GCCGAGAGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGG 240
DB 261 GCCGAGAGCGCTGCGCTTACCGGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGG 319
QY 241 GAGCGCTTGTGCCACCGCACTCTAGCGCGCGCGCGCGCGCTGCGCTGCGCTGCGG 300

```

```

DB 320 SAGGCGCTTCTGCCACCGCAACTSCAGCGCGCGCGCGCGCGCGCGCGCGCGCTC 379
QY 301 GCAGCTATCCGAGCTGCGCTCTTCGGGGGCGCTGCTGCGCGCGCGCGCGCTCCTCAG 360
DB 380 GCAGCNWCCGAGCTGCGCTCTWGGGGGCGCTGCTGCGCGCGCGCGCGCTCCTCAG 439
QY 361 CGTGCACGACGAGGCTGCGCTTCCGCGCTTCCGCGCTTCCGCGCTTCCGCGCTTCCG 420
DB 440 CGTGCACGACGAGGCTGCGCTTCCGCGCTTCCGCGCTTCCGCGCTTCCGCGCTTCCG 499
QY 421 GACTTCCGAGCGCGCGCGCGCTTCCGCGCTTCCGCGCTTCCGCGCTTCCGCGCTTCCG 480
DB 500 GACTTCCGAGCGCGCGCGCGCTTCCGCGCTTCCGCGCTTCCGCGCTTCCGCGCTTCCG 559
QY 481 CTCCTCCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
DB 560 CTCCTCCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 619
QY 541 ATGAGAGAGGACATGCGCATATATAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
DB 620 ATGAGAGAGGACATGCGCATATATAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 679
QY 601 CTGGAACCAAGTCAATATGAAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
DB 680 CTGGAACCAAGTCAATATGAAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 739
QY 661 AACTGGAAGACATCCATCACAGACATGAGAGTGGCGCTTCCGCGCTTCTTCAAGCGCTTT 720
DB 740 AACTGGAAGACATCCATCACAGACATGAGAGTGGCGCTTCCGCGCTTCTTCAAGCGCTTT 799
QY 721 TACGAGTCTCTGCGAGCTGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
DB 800 TACGAGTCTCTGCGAGCTGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 859
QY 781 CTTTCCATAGCAGATCAATATGTAAGTCTTGAAGTCTTGAAGTCTTGAAGTCTTGAAGTCT 840
DB 860 CTTTCCATAGCAGATCAATATGTAAGTCTTGAAGTCTTGAAGTCTTGAAGTCTTGAAGTCT 919
QY 841 CTCACCCCGAGTATAGGAGG 860
DB 920 CTCACCCCGAGTATAGGAGG 939

RESULT 14
AL558186 1201 bp mRNA linear EST 31-MAY-2003
LOCUS AL558186 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
DEFINITION Homo sapiens cDNA clone CS0DJ001YN15 5-PRIME, mRNA sequence.
ACCESSION AL558186
VERSION AL558186.2 GI:31279985
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jesses, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 15, 2001 this sequence version replaced gi:12902466.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 2233.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DJ001CG08QP1&cluster=2233.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DJ001CG08QP1.

FEATURES
Location/Qualifiers

```

```
source
1. 1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ001Y15"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      279 a 319 c 313 g 228 t
ORIGIN
Query Match      69.3%; Score 834; DB 9; Length 1201;
Best Local Similarity 96.5%; Pred. No. 1.1e-187;
Matches 881; Conservative 20; Mismatches 7; Indels 5; Gaps 5;
QY 1 ATGGAGCGGGCGCGCGGGCGCGCGCGCTGCTAGCGCTGCTGCGTGGCGCTGCGTGGCGCTGCGG 60
Db 95 ATGGANCGRRGCGCGGCGCGCGCGCTGCTAGCGCTGCTGCGTGGCGCTGCGTGGCGCTGCGG 154
QY 61 CTGCGCGCGCGCGCGCGCGCGCAATACGAAGCTTACAGCTTCCGAGCTTCCACGGGACGAG 120
Db 155 CTGCGCGCGCGCGCGCGCGCGCAATACGAAGCTTACAGCTTCCGAGCTTCCACGGGACGAG 214
QY 121 CTGATGCGCGCTCGAGTGGCGCTACCGGACGCGCTGACAGTACAGCGGCGGACGAGTGG 180
Db 215 CTGATGCGCGCTCGAGTGGCGCTACCGGACGCGCTGACAGTACAGCGGCGGACGAGTGG 274
QY 181 GCCAGAGCGCTGCGCTACCTGAGATCAGCTGCGGCTGCGGCTGCTGCTGCGGCGAGCAGC 240
Db 275 GCCAGAGCGCTGCGCTACCTGAGATCAGCTGCGGCTGCGGCTGCTGCTGCGGCGAGCAGC 334
QY 241 GAGCGCTTCTGCCACCGCACTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTC 300
Db 335 GAGCGCTTCTGCCACCGCACTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTC 394
QY 301 GCCAGCTATCCGAGCTGCGCTCTTTCGGGGGCGCTGCTGCGGCGCGCGCGCGCGCGCGCTC 360
Db 395 GCCAGCTATCCGAGCTGCGCTCTTTCGGGGGCGCTGCTGCGGCGCGCGCGCGCGCGCGCTC 454
QY 361 CGCTGCAAGCAGGCGCTTCCAGCGCTTCCGCGAGTCCCGCGCGCGCGCGCGCGCGCGCGCTC 420
Db 455 CGCTGCAAGCAGGCGCTTCCAGCGCTTCCCGCGAGTCCCGCGCGCGCGCGCGCGCGCGCTC 514
QY 421 GACTTCCAGCGCGCGCGCGCGCGCGCTTCCAGCTTCCGCTTCTTCTTCTTCTTCTTCTTCTT 480
Db 515 GACTTCCAGCGCGCGCGCGCGCGCGCTTCCAGCTTCCGCTTCTTCTTCTTCTTCTTCTTCTT 574
QY 481 CTGCGCAAGCGATGCGCGCTGCTCAACCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 540
Db 575 CTGCGCAAGCGATGCGCGCTGCTCAACCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 634
QY 541 ATGAAGAGGAACATGGCATAT-TATAGAGCGCTG-CCTGGTCCGAGGACTACATTAAAG 598
Db 635 ATGAAGAGGAACATGGCATATATATATATATATATATATATATATATATATATATATAT 694
QY 599 ACCTGGAACCAAGTATGAAGCGCTGTTCTTCCGAGCAGTGGCGGCGATACACCGGTG 658
Db 695 ACCTGGAACCAAGTATGAAGCGCTGTTCTTCCGAGCAGTGGCGGCGATACACCGGTG 754
QY 659 AGAAGCTGGAGAACATCATCACAGATGAGTGGCGCTTCCGAGCTTCTTCTTCTTCTTCTTCT 718
Db 755 AGAAGCTGGAGAACATCATCACAGATGAGTGGCGCTTCCGAGCTTCTTCTTCTTCTTCTTCT 813
QY 719 TTATACGAGTGTCTCGAGCGCTGCGAGGGTTTCAGGGAGATCAAGGACTTCAAGGATTTCT 778
Db 814 TTATACGAGTGTCTCGAGCGCTGCGAGGGTTTCAGGGAGATCAAGGACTTCAAGGATTTCT 873
QY 779 ACCTTCCATAGCAGATCATTTATGTAGAAGTCTTGGAAATGCAAAATACAGTGTGAAGAGA 838
Db 874 ACCTTCCATAGCAGATCATTTATGTAGAAGTCTTGGAAATGCAAAATACAGTGTGAAGAGA 932
QY 839 ACCTCACCACCGATTATAGGAGGCTATCCGGTTGAGAAATTTGGCTACCATGTATCAT 898
Db 933 A-CTCMCCCGCGATTAGGAGGCTATCCGGTTGAGAAATTTGGCTACCATGTATCAT 991
QY 899 ACTTGCGAGTTTCG 911
Db 992 ACTKGAGKTTTCG 1004
RESULT 15
AL550966
LOCUS
DEFINITION
AL550966 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0DI065YP05 5-PRIME, mRNA sequence.
ACCESSION
AL550966
VERSION
AL550966.2 GI:31272783
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li W.B., Gruber,C., Jessee,J., and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 15, 2001 this sequence version replaced gi:12888460.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 2233.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DI065CH03QPI&cluster=2233.f. Contact :
Peng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DI065CH03QPI.
FEATURES
Location/Qualifiers
1..1152
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI065YP05"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      266 a 311 c 285 g 227 t 63 others
ORIGIN
Query Match      68.3%; Score 822.2; DB 9; Length 1152;
Best Local Similarity 96.5%; Pred. No. 6.7e-185;
Matches 898; Conservative 10; Mismatches 16; Indels 7; Gaps 7;
QY 1 ATGAGCGCGGGCGCGCGGGCGCGCGCGCTGCTAGCGCTGCTGCGTGGCGCTGCGG 60
Db 51 ATGAGCGCGGGCGCGCGGGCGCGCGCGCTGCTAGCGCTGCTGCGTGGCGCTGCGG 110
QY 61 CTGCGCGCGGGCGCGCGCGCGCAATACGAAGCTTCCGAGCTTCCACGGGACGAG 120
Db 111 CTGCGCGCGGGCGCGCGCGCGCAATACGAAGCTTCCGAGCTTCCACGGGACGAG 169
QY 121 CTGATGCGCGCTCGAGTGGCGCTACCGGACGCGCTGGAAGTACAGCGCGGAGCAGTGG 180
Db 170 CTGATGCGCGCTCGAGTGGCGCTACCGGACGCGCTGGAAGTACAGCGCGGAGCAGTGG 229
QY 181 GCCGAGAGCGCTKGCTACCTGGAGATCAGCGCTGCGGCTGCACCGCTTGTGCGCGCAGC 240
```

Db 230 GCGAAGCGTGGCTACTTGGATCAGCTACGGCTNCACCGCTTGCTCGCGACAG- 288
Qy 241 GAGGCTTCTGCGAAGCGGAACTGCGAGCGCGCGCGAGCGCGCGCGCGCGCGCTC 300
Db 289 SAGGCTTCTGCGAAGCGGAACTGCGAGCGCGCGCGAGCGCGCGCGCGCGCTC 348
Qy 301 GCGAGCTATCCGAGCTGCGCTCTTCCGGGGCTTCTGCGCGCGCGCGCGCTCAAG 360
Db 349 GCGAGCTATCCGAGCTGCGCTCTTCCGGGGCTTCTGCGCGCGCGCGCGCTCAAG 408
Qy 361 GCGTGAAGAGCGCTGCGAGCTTCCGCGAGTCCGCGAGCGCGCGCGCGCGCTGCGG 420
Db 409 CGCTGCAA-CAGGGCGTGCAGCGCTTCCGCGAGTCCGCGAGCGCGCGCGCTGCGG 467
Qy 421 GACTTCCAGCGCGCGAGCGCTTACAACTTCCGCGAGTTCGCTTACTTCAAGGCAAAAT 480
Db 468 GACTTCCAGCGCGCGAGCGCTTACAACTTCCGCGAGTTCGCTTACTTCAAGGCAAAAT 527
Qy 481 CTCGCCAAGCGCTGCGCGCTGCTCAGCTTCTTACTGAGGATCTCTGATGAGGAAATG 540
Db 528 CTCGCCAAGCGCTGCGCGCTGCTCAGCTTCTTACTGAGGATCTCTGATGAGGAAATG 587
Qy 541 ATGAGAGGAAATGCGATGATAT-AAGAGCGCTGCTGCGCGAGGACTACATTAAAGA 599
Db 588 ATGAGAGGAAATGCGATGATAT-AAGAGCGCTGCTGCGCGAGGACTACATTAAAGA 647
Qy 600 CTGGAACCAAGTCAATGAAAGCTGTTCAATCCGAGCGAGTGGCGGATACAAAGGCTG 659
Db 648 CTGGAACCAAGTCAATGAAAGCTGTTCAATCCGAGCGAGTGGCGGATACAAAGGCTG 707
Qy 660 GAATGAGGAAATGCGATGATGAGGCTGCGCGAGTGGCGGCTTCCCGACTTCTTCAAAGGCTT 719
Db 708 GAATGAGGAAATGCGATGATGAGGCTGCGCGAGTGGCGGCTTCCCGACTTCTTCAAAGGCTT 767
Qy 720 TTACGAGTCTCGCGAGCTGCGAGGTTCCAGGAGATCAAGGACTTCAAGGATTTCTA 779
Db 768 TTACGAGTCTCGCGAGCTGCGAGGTTCCAGGAGATCAAGGACTTCAAGGATTTCTA 827
Qy 780 CCTTCCATAGCAGATCATATGATGAGTCTTGAATGCAAAATACAGTGTGAAGAGAA 839
Db 828 CCTTCCATAGCAGATCATATGATGAGTCTTGAATGCAAAATACAGTGTGAAGAGAA 886
Qy 840 CTCACCCCGAGTTATAGGAGGCTATCCGCTGAGAAATTTGCGTACCATGTATCATTA 899
Db 887 -CTCCACCCCGAGTTATAGGAGGCTATCCGCTGAGAAATTTGCGTACCATGTATCATTA 945
Qy 900 CTGCGAGTTTGCCTATTATTAAGTTGAACGAC 930
Db 946 CTGCGAGTTTG-CTATATAGTTGAMSRCC 975

Search completed: January 30, 2004, 18:26:30
Job time : 2528.98 secs

GenCore version 5.1.6
Copyright (C) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2004, 05:50:18 ; Search time 312.764 Seconds
(without alignments)
10382.969 Million cell updates/sec

Title: US-09-729-674-1_COPY_63_1265

Perfect score: 1203

Sequence: 1 atggagccggggcgccgggg.....tgggaactggagagaccagc 1203

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 13Jun03:*

```

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1202.2	99.9	3871	20 AAX60801	Human secreted protein
2	1202.2	99.9	3871	22 AAS59207	Human CDNA encodin
3	1202.2	99.9	3871	24 ABA90876	Human polynucleoti
4	1197.4	99.5	1961	22 AAH14525	Human CDNA sequenc
5	705	58.6	1586	21 AAH31138	Human colon cancer
6	656.6	54.6	818	22 AAH06808	Human CDNA clone
7	465.8	38.7	2347	24 AAL67662	Oesophagus cancer
8	406.4	33.8	408	21 AAH30585	Human colon cancer

9	156.4	13.0	1309	24 ABZ11344	Human polynucleoti
10	156.4	13.0	2152	24 ABZ11345	Human polynucleoti
11	156.4	13.0	2211	24 AAS17573	DNA encoding novel
12	156.4	13.0	2242	24 ABZ11346	Human polynucleoti
13	156.4	13.0	2524	22 AAH14513	Human CDNA sequenc
14	156.4	13.0	2563	22 AAF93800	Human CDNA encodin
15	156.4	13.0	2585	22 AAH14589	Human CDNA sequenc
16	156.4	13.0	2993	22 AAH16288	Human CDNA sequenc
17	154.8	12.9	2600	21 AAC64725	Human tumour suppr
18	154.8	12.9	2829	21 AAC64724	Human tumour suppr
19	152.6	12.7	2577	22 AAD08489	Human secreted pro
20	150.4	12.5	2615	22 AAD08507	Human secreted pro
21	147.4	12.3	2322	21 AAC64727	Mouse tumour suppr
22	147.4	12.3	2416	21 AAC64726	Mouse tumour suppr
23	130.8	10.9	2127	24 AAS17572	DNA encoding novel
24	126	10.5	618	21 AA280131	Human colon cancer
25	121.4	10.1	325	22 AA182020	Human polynucleoti
26	116.2	9.7	1026	24 ABQ26350	Oligonucleotide fo
27	116.2	9.7	1026	24 ABQ26351	Oligonucleotide fo
28	115.8	9.6	1025	24 ABQ27312	Oligonucleotide fo
29	115.8	9.6	1025	24 ABQ27313	Oligonucleotide fo
30	109.2	9.1	2665	25 ACA03965	CDNA downregulated
31	99.2	8.2	652	22 AAF93998	Primer specific fo
32	99.2	8.2	738	22 AAH06612	Human CDNA clone
33	99	8.2	703	22 AAH06792	Human CDNA clone
34	98.6	8.2	774	22 AAH06885	Human CDNA clone
35	98.6	8.2	800	22 AAH06448	Human CDNA clone
36	98.6	8.2	2789	22 AAH14459	Human CDNA sequenc
37	97.8	8.1	2753	24 ABK94910	Human novel polynu
38	97.8	8.1	2753	24 ABK94933	Human novel polynu
39	88	7.3	1025	24 ABQ27310	Oligonucleotide fo
40	88	7.3	1025	24 ABQ27311	Oligonucleotide fo
41	88	7.3	1026	24 ABQ26352	Oligonucleotide fo
42	88	7.3	1026	24 ABQ26353	Oligonucleotide fo
43	87	7.2	349	21 AAF18376	Oligonucleotide fo
44	73.2	6.1	550	25 ACA57060	Lung cancer associ
45	68.2	5.7	3897	24 ABL60615	Human adipocyte se
					Maize SBMu200 geno

ALIGNMENTS

RESULT 1

AAX60801

ID AAX60801 standard; DNA; 3871 BP.

XX

AC AAX60801;

XX

DT 09-AUG-1999 (first entry)

XX

XX Human secreted protein encoding DNA (clone bd306-7).

DE

XX Secreted protein; kidney; lung; brain; blood; testis; bone marrow;

KW nutritional activity; cytokine; cell proliferation; immune stimulation;

KW hematopoiesis regulation; tissue growth; thrombolytic; gene therapy;

KW anti-inflammatory; tumour invasion; ss.

XX

OS Homo sapiens.

XX

PN WO9926961-A1.

XX

PD 03-JUN-1999.

XX

PF 24-NOV-1998; 98WO-US25149.

XX

PR 23-NOV-1998; 98US-0197886.

PR

26-NOV-1997; 97US-0066804.

XX

PA (GEM) GENETICS INST INC.

XX

PI Agostino MJ, Clark HF, Collins-Racie LA, Evans C;

Reichel K, Jacobs K, Lavallie ER, McCoy JM, Merberg D;

Steininger RJ, Treacy M, Wong GG;

XX WFI; 1999-357809/30.
DR P-PSDB; AAY17219.
XX
PT New polynucleotides encoding secreted proteins
XX
PS Claim 1; Page 110-111; 133pp; English.
XX
CC The invention relates to secreted proteins (AAY17219-228) encoded by
CC polynucleotides obtained from human fetal kidney, adult lung, adult
CC kidney, adult brain, adult blood, adult testes, and fetal brain and
CC murine adult bone marrow cDNA libraries. The secreted protein nucleic
CC acid sequences (X6801-811) correspond to clones bd306-7, gJ283-6,
CC fk317-3, k213-2x, na316-1, nf93-20, np164-1, pe204-1, yal-1 and yb-1,
CC (all clones are deposited as ATCC 98599). The PNs and proteins are
CC predicted to have biological activities which would make them suitable
CC for treating, preventing or ameliorating medical conditions in humans and
CC animals, although no supporting data is given. Suggested activities
CC include nutritional activity, cytokine and cell
CC proliferation/differentiation activity, immune stimulating (e.g. as
CC vaccines) or suppressing activity, hematopoiesis regulating activity,
CC tissue growth activity, activin/inhibin activity,
CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,
CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumour
CC invasion suppressor activity, and tumour inhibition activity. The PNs are
CC also stated to be useful for gene therapy.

XX
SQ Sequence 3871 BP; 965 A; 913 C; 1033 G; 958 T; 2 other;
XX
Query Match 99.9%; Score 1202.2; DB 20; Length 3871;
Best Local Similarity 100.0%; Pred. No. 1.7e-253;
Matches 1203; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 ATGGAGCGGGCGCGGGGGCGCGCGCTGCTAGCGTGTGTGCGTGGCGTGGCGG 60
DB 63 ATGGAGCGGGCGCGGGGGCGCGCGCTGCTAGCGTGTGTGCGTGGCGTGGCGG 122
QY 61 CTGCGCGCGGGCGCGCCCAATAGCAACGCTACAGCTTCGCGAGCTTCCACGGGACGAG 120
DB 123 CTGCGCGCGGGCGCGCCCAATAGCAACGCTACAGCTTCCGCGAGCTTCCACGGGACGAG 182
QY 121 CTGATGCGCGCTGAGTGGCGCTACGGGACGCGCTGACAGTACAGCGCGGACGACTGG 180
DB 183 CTGATGCGCGCTGAGTGGCGCTACGGGACGCGCTGACAGTACAGCGCGGACGACTGG 242
QY 181 GCGGAGAGCSTKGGCTACTCTGGAGATCAGCGTGGCGGCTGACCGCTTGTGCGCGACAGC 240
DB 243 GCGGAGAGCSTKGGCTACTCTGGAGATCAGCGTGGCGGCTGACCGCTTGTGCGCGACAGC 302
QY 241 GAGGCGTTTGCACCGCAACTGACAGCGCGCGCGCGAGCGCGGAGCGCGCGCGCGCTTC 300
DB 303 GAGGCGTTTGCACCGCAACTGACAGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCTTC 362
QY 301 GCGAGCTATCCGAGCGTGGCGCTTCTCGGGGGCGCTGTGCGCGCGCGCGCGCTTCCTCAAG 360
DB 363 GCGAGCTATCCGAGCGTGGCGCTTCTCGGGGGCGCTGTGCGCGCGCGCGCTTCCTCAAG 422
QY 361 CGGTGCAAGCAGGGCGTGGCAGCGCTTCGCGCAGTCCAGCGCCAGCGCGCGAGGTGCTGGCG 420
DB 423 CGGTGCAAGCAGGGCGTGGCAGCGCTTCGCGCAGTCCAGCGCCAGCGCGCGAGGTGCTGGCG 482
QY 421 GACTTCCAGCGCGCGAGCGCTTACAAGTTCCTGCAAGTTCGCTACTTCAAGGCAATAAT 480
DB 483 GACTTCCAGCGCGCGAGCGCTTACAAGTTCCTGCAAGTTCGCTACTTCAAGGCAATAAT 542
QY 481 CTCCCCAAGCCATGCGCGTGTCTACACCTTTCTACTGAGCATCTCTGATGACGAATG 540
DB 543 CTCCCCAAGCCATGCGCGTGTCTACACCTTTCTACTGAGCATCTCTGATGACGAATG 602
QY 541 ATGAAGAGGAAATGCGCATATTATTAAGAGCGCTGCGTGGTGGCGGAGCACTACATTAAGAC 600
DB 603 ATGAAGAGGAAATGCGCATATTATTAAGAGCGCTGCGTGGTGGCGGAGCACTACATTAAGAC 662

PD 11-OCT-2001.
XX 22-MAR-2001; 2001WO-US09369.
PF 30-MAR-2000; 2000US-0539330.
PR 04-DEC-2000; 2000US-0729674.
XX (GEMV) GENETICS INST INC.
PA
XX
PI Jacobs K, McCoy JM, Lavallie E, Collins-racie LA, Evans C;
PI Treacy M, Agostino MJ, Steininger RJ, Spaulding V, Wong GG;
PI Clark H, Fichtel K, Merberg D;
DR
XX
DR WPI; 2001-639363/73.
XX P-PSDB; AAU38989.
XX
PT Secreted human proteins, useful as vaccine for treating various
PT diseases such as autoimmune disorders (e.g. multiple sclerosis), and
PT nervous system disorders (e.g. stroke) -
XX
PS Claim 1; Page 456-457; 619pp; English.
XX
CC The invention relates to novel human secreted proteins, the nucleic
CC acids encoding them. The protein may exhibit cytokine, cell proliferation
CC or cell differentiation activity or may induce production of other
CC cytokines in certain cell populations and may exhibit immune stimulating
CC or immune suppressing activity, which is useful for the treatment of
CC various immune deficiencies and disorders e.g. severe combined
CC immunodeficiency (SCID), autoimmune disorders e.g. multiple sclerosis,
CC systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary
CC inflammation. The proteins are also useful in the treatment of diseases
CC and disorders including tissue, skin and organ transplantation and in
CC graft-versus-host diseases (GVHD), in the induction of tumour immunity,
CC myeloid or lymphoid cell deficiencies, wound healing and tissue repair,
CC in the treatment of burns, incisions and ulcers; as well as in treatment
CC of periodontal disease, osteoporosis or osteoarthritis, mediated by
CC inflammatory processes, diseases of the peripheral nervous system,
CC Alzheimer's, Parkinson's disease, Huntington's disease,
CC amyotrophic lateral sclerosis, and Shy-Drager syndrome, infections,
CC infarction of cardiac and central nervous system vessel e.g. stroke,
CC sepsis, inflammatory bowel disease, ulcers, bone regeneration. The
CC protein, having activin- or inhibin-related activities is useful as a
CC contraceptive based on the ability of inhibin to decrease fertility in
CC female mammals and decrease spermatogenesis in male mammals. The
CC proteins and nucleic acids are also useful as food supplements. The
CC present sequence encodes a secreted protein of the invention.
XX
SQ Sequence 3871 BP; 965 A; 913 C; 1033 G; 958 T; 2 other;

Query Match 99.9%; Score 1202.2; DB 22; Length 3871;
Best Local Similarity 100.0%; Pred. No. 1.7e-253;
Matches 1203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCGGGGGCGCGGGGGCGCGCGCGCTGCTAGCGCTGCTGCGGCTCGCGGCTCGCGG 60
DB 63 ATGAGCGGGGGCGCGGGGGCGCGCGCGCTGCTAGCGCTGCTGCGGCTCGCGGCTCGCGG 122
QY 61 CTGCGCGCGGGCGAG 120
DB 123 CTGCGCGCGGGGGCGAG 182
QY 121 CTGATGCGGCTCGAGTCGGGCTACCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 180
DB 183 CTGATGCGGCTCGAGTCGGGCTACCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 242
QY 181 GCGGAGCGCTKGCTACCTGGAGATCAGCTCGGCTGCGCGCGCGCGCGCGCGCGCGCGAG 240
DB 243 GCGGAGCGCTKGCTACCTGGAGATCAGCTCGGCTGCGCGCGCGCGCGCGCGCGCGCGAG 302
QY 241 GAGCGCTTCTGCTC 300
DB 303 GAGCGCTTCTGCTC 362

RESULT 3

ABA90876
ID ABA90876 standard; cDNA; 3871 BP.
XX
AC ABA90876;
XX
DT 14-FEB-2002 (first entry)
XX

QY 301 GCCAGCTATCCGAGCTGCGGCTCTTTCGGGGGCGCTGCTGCGGCGCGCGCACTGCGCTCAAG 360
DB 363 GCCAGCTATCCGAGCTGCGGCTCTTTCGGGGGCGCTGCTGCGGCGCGCGCACTGCGCTCAAG 422
QY 361 CGCTGCAAGCAGGCGCTGCGGCTTCCGCGAGTCCGCGAGTCCGCGAGCGCGAGGCTGCGGCG 420
DB 423 CGCTGCAAGCAGGCGCTGCGGCTTCCGCGAGTCCGCGAGTCCGCGAGCGCGAGGCTGCGGCG 482
QY 421 GACTTCCAGCGCGCGAGCGCTTACAGTTCCTGCGAGTTCCTGCGTTCCTTCAAGGCAATAT 480
DB 483 GACTTCCAGCGCGCGAGCGCTTACAGTTCCTGCGAGTTCCTGCGTTCCTTCAAGGCAATAT 542
QY 481 CTCCCAAGAGCCATCGCGCTGCTCAGCTTCTTACTGAAAGCATCTCTGATCAAGCAATAT 540
DB 543 CTCCCAAGAGCCATCGCGCTGCTCAGCTTCTTACTGAAAGCATCTCTGATCAAGCAATAT 602
QY 541 ATGAAGAGCAATGCGCATATTAAGAGCGCTGCTGCGAGGAGTCAAGGATTCCTTCAAGG 600
DB 603 ATGAAGAGCAATGCGCATATTAAGAGCGCTGCTGCGAGGAGTCAAGGATTCCTTCAAGG 662
QY 601 CTGGAACCAAGTCAATGAAAGCGCTTTCATCCGAGCAGTCCGGGCGCATACACGCTGAG 660
DB 663 CTGGAACCAAGTCAATGAAAGCGCTTTCATCCGAGCAGTCCGGGCGCATACACGCTGAG 722
QY 661 AACTGGAGAACATCCATCAGCATGAGCTGCGGCTTCCGCGAGTTCCTTCAAGGCTTTC 720
DB 723 AACTGGAGAACATCCATCAGCATGAGCTGCGGCTTCCGCGAGTTCCTTCAAGGCTTTC 782
QY 721 TAGAGTGTCTCGAGCGCTGCGAGGCTTCCGAGGAGTCAAGGATTCCTTCAAGGATTCCTAC 780
DB 783 TAGAGTGTCTCGAGCGCTGCGAGGCTTCCGAGGAGTCAAGGATTCCTTCAAGGATTCCTAC 842
QY 781 CTTTCCATAGCAGATCATATGTAAGTTCGGAATGCAAAATACAGTGTGAAGAGAAC 840
DB 843 CTTTCCATAGCAGATCATATGTAAGTTCGGAATGCAAAATACAGTGTGAAGAGAAC 902
QY 841 CTCACCCAGTTATAGGAGGCTATCCGCTTGAGAAATTTGCGCTACCATGTATCATTTAC 900
DB 903 CTCACCCAGTTATAGGAGGCTATCCGCTTGAGAAATTTGCGCTACCATGTATCATTTAC 962
QY 901 TTGCAATTTGCTTATTAAGTTCGAACTGAAAGTCAAGGAGTCCGCGCGCTGCGAGTCAAG 960
DB 963 TTGCAATTTGCTTATTAAGTTCGAACTGAAAGTCAAGGAGTCCGCGCGCTGCGAGTCAAG 1022
QY 961 TATCTGCTTTTATGACAGATGCAAGTTCATGAGCAGACCTGCTGCTGATTTACAGTAC 1020
DB 1023 TATCTGCTTTTATGACAGATGCAAGTTCATGAGCAGACCTGCTGCTGATTTACAGTAC 1082
QY 1021 CACAGGACACTTGGGGCTCTCGGATGAGCACTTCCAGCGCGAGACCTGAGCAGTTCAG 1080
DB 1083 CACAGGACACTTGGGGCTCTCGGATGAGCACTTCCAGCGCGAGACCTGAGCAGTTCAG 1142
QY 1081 TTCTTTATGTCACACACTCCAGAGGAGTGTATGACCTTGTGTAAGGAAATATATATG 1140
DB 1143 TTCTTTATGTCACACACTCCAGAGGAGTGTATGACCTTGTGTAAGGAAATATATATG 1202
QY 1141 GATCATGATGAGGAGAGTGTGGAATATGATGATGACCTTCTTGGAACTGGAGGAGAC 1200
DB 1203 GATCATGATGAGGAGAGTGTGGAATATGATGATGACCTTCTTGGAACTGGAGGAGAC 1262
QY 1201 AGC 1203
DB 1263 AGC 1265

DE Human polynucleotide SEQ ID NO 1.
 XX Human; clone bd306-7; clone yb8-1; ATCC number 98599; gene therapy;
 KW immune disorder; bacterial infection; fungal infection; cancer; tumour;
 KW autoimmune disorder; systemic lupus erythematosus; wound; ulcer; inhibitor;
 KW osteoarthritis; osteoarthritis; nervous system disorder; neuropathy;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease; activin;
 KW haemophilia; cardiac infarction; stroke; sepsis; arthritis; vulnerability;
 KW ischaemia-reperfusion injury; inflammatory bowel disease; chemotactic;
 KW Crohn's disease; cytostatic; anti-inflammatory; immunomodulator;
 KW neuroprotective; haemostatic; thrombolytic; anti-inflammatory; ss.
 XX Homo sapiens.
 OS
 XX
 PN US2001039335-A1.
 XX
 PD 08-NOV-2001.
 XX
 PF 04-DEC-2000; 2000US-0729674.
 XX
 PR 10-APR-1997; 97US-126425P.
 PR 04-DEC-1997; 97US-067454P.
 PR 20-DEC-1997; 97US-068379P.
 PR 02-JAN-1998; 98US-070346P.
 PR 07-JAN-1998; 98US-070643P.
 PR 08-JAN-1998; 98US-070755P.
 PR 13-JAN-1998; 98US-071304P.
 PR 22-JAN-1998; 98US-072113P.
 PR 30-JAN-1998; 98US-073095P.
 PR 18-FEB-1998; 98US-075038P.
 PR 30-MAR-2000; 2000US-0539330.
 PR 23-NOV-1998; 98US-0197886.
 XX
 PA (JACO/) JACOBS K.
 PA (MCCO/) MCCOY J M.
 PA (LAVA/) LAVALLIE E R.
 PA (COLL/) COLLINS-RACIE L A.
 PA (EVAN/) EVANS C.
 PA (MERB/) MERBERG D.
 PA (TREAR/) TREACY M.
 PA (AGOS/) AGOSTINO M J.
 PA (STEI/) STEININGER R J.
 PA (SPAU/) SPAULDING V.
 PA (WONG/) WONG G G.
 PA (CLAR/) CLARK H.
 PA (FECH/) FECHTEL K.
 XX
 PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
 PI Merberg D, Treacy M, Agostino MJ, Steininger RJ, Spaulding V;
 PI Wong GG, Clark H, Fechtel K;
 XX
 DR WPI: 2002-040725/05.
 DR P-PSDB; ABB55698.
 XX
 PT New secreted proteins and encoding polynucleotides, useful in gene
 PT therapies, particularly for preventing or treating autoimmune
 PT disorders, cancer, graft-versus-host disease, wound, osteoporosis,
 PT stroke or inflammations
 XX
 PS Claim 1; Page 175-177; 349pp; English.
 XX
 CC The invention relates to isolated polynucleotides (ABA90876-ABA90568 and
 CC ABA90980) and encoded proteins (ABB55698-ABB55900), especially
 CC polynucleotides SEQ ID NO 1 (ABA90876) and SEQ ID NO 19 (ABA90885) and
 CC proteins SEQ ID NO 2 (ABB55698) and SEQ ID NO 20 (ABB55707) contained in
 CC clones bd306-7 and yb8-1 respectively and the clones bd306-7 and yb8-1
 CC are deposited with the American Type Culture Collection (ATCC) with
 CC accession number 98599. The polynucleotides and encoded polypeptides have
 CC cytosolic, anti-inflammatory, immunomodulator, vulnerability,
 CC neuroprotective, activin, inhibitor, chemotactic, haemostatic, thrombolytic
 CC and anti-inflammatory activity and acting as cytokine modulators,
 CC haematopoiesis regulators, tissue growth modulators and/or cadherin
 CC suppressors. The polypeptides and polynucleotides are useful in gene

therapies, particularly for preventing, treating or ameliorating any of
 the following diseases: immune deficiency and disorders; e.g. bacterial
 or fungal infections, autoimmune disorders, cancer, systemic lupus
 erythematosus or graft-versus-host disease, myeloid or lymphoid cell
 deficiencies; wound, burns, incisions and ulcers, osteoporosis or
 osteoarthritis; central and peripheral nervous system diseases or
 neuropathies, e.g. Alzheimer's, Parkinson's disease, Huntington's
 disease, amyotrophic lateral sclerosis or Shy-Drager syndrome;
 haemophilia, cardiac infarction or stroke; inflammations, shock, sepsis
 or systemic inflammatory response syndrome; ischaemia-reperfusion
 injury, endotoxin lethality, arthritis, inflammatory bowel disease or
 Crohn's disease; or tumours or cancers, pemphigus vulgaris or pemphigus
 foliaceus.
 XX
 XX
 Sequence 3871 BP; 965 A; 913 C; 1033 G; 958 T; 2 other;
 Query Match 99.98; Score 1202.2; DB 24; Length 3871;
 Best Local Similarity 100.0%; Pred. No. 1.7e-253;
 Matches 1203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGGAGCCGGGGCGCGGGGGGGCGGGCGCTCTAGCGCTGCTGCGTGGCGCTCGCG 60
 Db
 QY 63 ATGGAGCCGGGGCGCGGGGGGGCGGGCGCTCTAGCGCTGCTGCGTGGCGCTCGCG 122
 Db
 QY 61 CTGCGCGCGGGCGCGGGGGGGCGGGCGCTCTAGCGCTGCTGCGTGGCGCTCGCG 120
 Db
 QY 123 CTGCGCGCGGGCGCGGGGGGGCGGGCGCTCTAGCGCTGCTGCGTGGCGCTCGCG 182
 Db
 QY 121 CTGATGCCGCTCGAGTCGGGCTTACCGGCGACGCGCTGCAAGTACAGCGCGGAGCACTGG 180
 Db
 QY 183 CTGATGCCGCTCGAGTCGGGCTTACCGGCGACGCGCTGCAAGTACAGCGCGGAGCACTGG 242
 Db
 QY 181 GCGGAGAGCSTKGCTACCTGGAGATCAGCTCGGCTGCGCTGCTGCGTGGCGCAGC 240
 Db
 QY 243 GCGGAGAGCSTKGCTACCTGGAGATCAGCTCGGCTGCGCTGCTGCGTGGCGCAGC 302
 Db
 QY 241 GAGGCGCTTCTGCCACCGCAACTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCTC 300
 Db
 QY 303 GAGGCGCTTCTGCCACCGCAACTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCTC 362
 Db
 QY 301 GCAGCTATCCGAGCTGCGCTCTTTCGGGGGCGCTGCTGCGGGCGCGCACTGCTCAAG 360
 Db
 QY 363 GCAGCTATCCGAGCTGCGCTCTTTCGGGGGCGCTGCTGCGGGCGCGCACTGCTCAAG 422
 Db
 QY 361 CGTGTCAAGCAGGCGCTGCGCGCTTCCGCGAGTCCGCGCGCGCGCGCGCGCGCGCG 420
 Db
 QY 423 CGTGTCAAGCAGGCGCTGCGCGCTTCCGCGAGTCCGCGCGCGCGCGCGCGCGCG 482
 Db
 QY 421 GACTTCCAGCGCGCGCGCGCTTACAGGTTCTTGCAGTTCGCTTACTTCAAGGCAATAT 480
 Db
 QY 483 GACTTCCAGCGCGCGCGCGCTTACAGGTTCTTGCAGTTCGCTTACTTCAAGGCAATAT 542
 Db
 QY 481 CTCGCCAAGAGCCATCGCGCTGCTCACACCTTTCTTACTGAAGCATCTGATGACGAATG 540
 Db
 QY 543 CTCGCCAAGAGCCATCGCGCTGCTCACACCTTTCTTACTGAAGCATCTGATGACGAATG 602
 Db
 QY 541 ATGAAGAGGAACATGCGCATATTTAAGAGCGCTGCTGCGCGCGCGCGCGCGCGCG 600
 Db
 QY 603 ATGAAGAGGAACATGCGCATATTTAAGAGCGCTGCTGCGCGCGCGCGCGCGCGCG 662
 Db
 QY 601 CTGGAACCAAGTCATATGAAGAGCCCTTTCATCCGAGCAGTTCGCGGCGCATACAACGGTGA 660
 Db
 QY 663 CTGGAACCAAGTCATATGAAGAGCCCTTTCATCCGAGCAGTTCGCGGCGCATACAACGGTGA 722
 Db
 QY 661 AACTGGAGAACATCCATCACAGATGAGTGGGCGCTTCCGCGCTTCTTCAAGAGCCCTT 720
 Db
 QY 723 AACTGGAGAACATCCATCACAGATGAGTGGGCGCTTCCGCGCTTCTTCAAGAGCCCTT 782
 Db
 QY 721 TACGAGTGTCTCGAGCTGCGAGGTTTCAGGGAGATCAAGGACTTCAAGGATTTTAC 780
 Db
 QY 783 TACGAGTGTCTCGAGCTGCGAGGTTTCAGGGAGATCAAGGACTTCAAGGATTTTAC 842
 Db
 QY 781 CTTTCCATAGCAGATCATATTATGTAGAGTTCTGGATTCGGAATCAAAATACAGTGTGAAGAGAC 840

Db 843 CTTTCCATAGCAGATCATATATAGAGTTCTGGAATGCAAAATACAGTGTGAAGAGAAC 902
QY 841 CTCACCCAGTTATAGGAGCTATCCGGTTGAGAAATTTGGCTACCATGTATCATATAC 900
Db 903 CTCACCCAGTTATAGGAGCTATCCGGTTGAGAAATTTGGCTACCATGTATCATATAC 962
QY 901 TTGCAGTTGGCTATATTAAGTTGAACGACCTGAAGATGAGCCCTGTGACGTGAGC 960
Db 963 TTGCAGTTGGCTATATTAAGTTGAACGACCTGAAGATGAGCCCTGTGACGTGAGC 1022
QY 961 TATCTGCTCTTTGATCAGAAATGACAGGTTCATGAGCAGACCTGCTGTTATACAGTAC 1020
Db 1023 TATCTGCTCTTTGATCAGAAATGACAGGTTCATGAGCAGACCTGCTGTTATACAGTAC 1082
QY 1021 CACAGGACACTTGGGGCTCTCGATGAGCACTTCCAGCCAGACCTGAAGCAGTTTCAG 1080
Db 1083 CACAGGACACTTGGGGCTCTCGATGAGCACTTCCAGCCAGACCTGAAGCAGTTTCAG 1142
QY 1081 TTCTTTAATGTGACCACTCCAGAGGAGCTGTATGACTTTGCTAAGGAAATATATATG 1140
Db 1143 TTCTTTAATGTGACCACTCCAGAGGAGCTGTATGACTTTGCTAAGGAAATATATATG 1202
QY 1141 GATGATGATGAGGAGAGTTGTGGAATATATGATGATGACCTTTGGAACCTGGAGGAGC 1200
Db 1203 GATGATGATGAGGAGAGTTGTGGAATATATGATGATGACCTTTGGAACCTGGAGGAGC 1262
QY 1201 AGC 1203
Db 1263 AGC 1265

RESULT 4

AAH14525
ID AAH14525 standard; cDNA; 1961 BP.
XX AAH14525;
AC
XX
XX 26-JUN-2001 (first entry)
XX Human cDNA sequence SEQ ID NO:12066.
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX Homo sapiens.
XX
XX EP1074617-A2.
XX
XX 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-0116126.
XX
XX 29-JUL-1999; 99JP-0248036.
XX 27-AUG-1999; 99JP-0300253.
XX 11-JAN-2000; 2000JP-0118776.
XX 02-MAY-2000; 2000JP-0183767.
XX 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayaashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX

Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs -

Claim 8; SEQ ID 12066; 2537pp + CD ROM; English.

The present invention describes primer sets for synthesizing 5602

full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo- α primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH95893 represent human amino acid sequences; AAH92446 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

Sequence 1961 BP; 417 A; 579 C; 481 G; 484 T; 0 other;

Query Match 99.5%; Score 1197.4; DB 22; Length 1961;

Best Local Similarity 99.6%; Pred. No. 1.6e-252;

Matches 1198; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGAGAGCGGGCGCGGGCGCGCGCTGTAGCGCTGTGCTGTGCTGTGCTGTGCGG 60
Db 40 ATGAGAGCGGGCGCGGGCGCGCGCTGTAGCGCTGTGCTGTGCTGTGCTGTGCGG 99
QY 61 CTGCGCGCGGGCGCGGGCGCGCGCTGTAGCGCTGTGCTGTGCTGTGCTGTGCGG 120
Db 100 CTGCGCGCGGGCGCGGGCGCGCGCTGTAGCGCTGTGCTGTGCTGTGCTGTGCGG 159
QY 121 CTGATGCGCGCTGAGTGGCGCTACCGGCAAGCTGGAAGTACAGCGGCGGAGCTGG 180
Db 160 CTGATGCGCGCTGAGTGGCGCTACCGGCAAGCTGGAAGTACAGCGGCGGAGCTGG 219
QY 181 GCCGAGAGCTKGGCTACTTGGAGATCAGCTTGGGGCTGCAACCGCTTCTCGCGGACAGC 240
Db 220 GCCGAGAGCTKGGCTACTTGGAGATCAGCTTGGGGCTGCAACCGCTTCTCGCGGACAGC 279
QY 241 GAGGCTTTCTGCCACCGCACTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTC 300
Db 280 GAGGCTTTCTGCCACCGCACTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTC 339
QY 301 GCCAGCTATCCCGAGCTGCGCTCTTTCGGGGGCGCTGTGCGCGCGCGCGCGCTCAAG 360
Db 340 GCCAGCTATCCCGAGCTGCGCTCTTTCGGGGGCGCTGTGCGCGCGCGCGCTCAAG 399
QY 361 CGCTGCAAGAGCGGCTGCGAGCGCTTCCGAGTCCCGAGCGCGCGCGCGCGCGCGCTC 420
Db 400 CGCTGCAAGAGCGGCTGCGAGCGCTTCCGAGTCCCGAGCGCGCGCGCGCGCGCTC 459
QY 421 GACTTCCAGCGCGCGAGCGCTCAAGTTCCTGAGTTCGCTTACTTCAAGGCGAAATAT 480
Db 460 GACTTCCAGCGCGCGAGCGCTCAAGTTCCTGAGTTCGCTTACTTCAAGGCGAAATAT 519
QY 481 CTCGCCAAAGCCATGCGCGCTGTACACCTTTCTACTGAAGCATCTGTATGAGCAATG 540
Db 520 CTCGCCAAAGCCATGCGCGCTGTACACCTTTCTACTGAAGCATCTGTATGAGCAATG 579
QY 541 ATGAAGAGGACATGGCATATTATAGAGCTTGGCTGTGCGAGCGCGCGCGCGCGCTC 600
Db 580 ATGAAGAGGACATGGCATATTATAGAGCTTGGCTGTGCGAGCGCGCGCGCGCGCTC 639
QY 601 CTGGAACCAAGTCATATGAAGCTTTTCATCCGAGCAGTGGGGGCGATACACCGGTGAG 660
Db 640 CTGGAACCAAGTCATATGAAGCTTTTCATCCGAGCAGTGGGGGCGATACACCGGTGAG 699
QY 661 AACTGGAGAACATCCATCACAGATGGAGCTGGCGCTTCCCGACTTCTTCAAGCGCTTT 720

QY 1097 CACTCAGAGGAGCTGTATGACTTCTTAAGGAAATATAATGATGATGAGGAG 1156
Db 601 CACTCAGAGGAGCTGTATGACTTCTTAAGGAAATATAATGATGATGAGGAG 660
QY 1157 AAGTTGTGGAATATGATGATGACCTCTTGGAACTGGAGGAGACCAAGC 1203
Db 661 AAGTTGTGGAATATGATGATGACCTCTTGGAACTGGAGGAGACCAAGC 707

RESULT 6

AAH06808

ID AAH06808 standard; cDNA; 818 BP.

AC AAH06808;

XX 26-JUN-2001 (first entry)

XX Human cDNA clone (5'-primer) SEQ ID NO:3643.

DE Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

XX 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602

XX full-length cDNAs defined in the specification, and for the detection

XX and/or diagnosis of the abnormality of the proteins encoded by the

XX full-length cDNAs -

XX Claim 1; SEQ ID 3643; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602

XX full-length cDNAs defined in the specification. Where a primer set

XX comprises: (a) an oligo-dr primer and an oligonucleotide complementary

XX to the complementary strand of a polynucleotide which comprises one of

XX the 5602 nucleotide sequences defined in the specification, where the

XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination

XX of an oligonucleotide comprising a sequence complementary to the

XX complementary strand of a polynucleotide which comprises a 5'-end

XX sequence and an oligonucleotide comprising a sequence complementary to a

XX polynucleotide which comprises a 3'-end sequence, where the

XX oligonucleotide comprises at least 15 nucleotides and the combination of

XX the 5'-end sequence/3'-end sequence is selected from those defined in

XX the specification. The primer sets can be used in antisense therapy and

XX in gene therapy. The primers are useful for synthesizing polynucleotides,

XX particularly full-length cDNAs. The primers are also useful for the

XX detection and/or diagnosis of the abnormality of the proteins encoded by

XX the full-length cDNAs. The primers allow obtaining of the full-length

XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and

XX AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to

XX AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632

XX represent oligonucleotides, all of which are used in the exemplification

XX of the present invention.

XX Sequence 818 BP; 159 A; 271 C; 236 G; 148 T; 4 other;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Query Match 54.6%; Score 656.6; DB 22; Length 818;
Best Local Similarity 96.6%; Pred. No. 3.1e-134;
Matches 700; Conservative 2; Mismatches 18; Indels 5; Gaps 3;

QY 1 ATGGAGCCGGGGCGCGGGGGCGCGGGCGCTGTAGCGCTGTGTGCGTGTGCGCTCGCG 60
Db 40 ATGGAGCCGGGGCGCGGGGGCGCGGGCGCTGTAGCGCTGTGTGCGTGTGCGCTCGCG 99
QY 61 CTGCGCGCGGGCGCGCGCGCAATACGAACCTACAGCTTCGGGAGCTTCCACCGGACGAG 120
Db 100 CTGCGCGCGGGCGCGCGCGCAATACGAACCTACAGCTTCGGGAGCTTCCACCGGACGAG 159
QY 121 CTGATCCCGCTCGAGTTCGCGCTTACCGGCAACGCGCTGGACAAGTACAGCGCGGACACTGG 180
Db 160 CTGATCCCGCTCGAGTTCGCGCTTACCGGCAACGCGCTGGACAAGTACAGCGCGGACACTGG 219
QY 181 GCGGAGAGGCTGGGCTACCTGGAGATCAGCTTGGCGGTGACCGCTTGTGTGCGGACAGC 240
Db 220 GCGGAGAGGCTGGGCTACCTGGAGATCAGCTTGGCGGTGACCGCTTGTGTGCGGACAGC 279
QY 241 GAGGCTTCTGCCACCGCAACTGCAGCGCGCGCGCGAGCCGCGAGCCGCGCGCGCTC 300
Db 280 GAGGCTTCTGCCACCGCAACTGCAGCGCGCGCGCGAGCCGCGAGCCGCGCGCGCTC 339
QY 301 GCGAGCTATCCGAGCTGCGGCTCTTGGGGGCGCTGTGTGCGCGCGCGCACTGCTCAAG 360
Db 340 GCGAGCTATCCGAGCTGCGGCTCTTGGGGGCGCTGTGTGCGCGCGCGCACTGCTCAAG 399
QY 361 CGCTGCAAGCAGGGGCTGCGAGCCTTCCGCCAGTCCCGAGCCCGAGCCGCGCGCGCTGCTGGG 420
Db 400 CGCTGCAAGCAGGGGCTGCGAGCCTTCCGCCAGTCCCGAGCCCGAGCCGCGCGCGCTGCTGGG 459
QY 421 GACTTCCAGCGCGCGGAGCCCTTACAAAGTTCTCGAGTTCGCTTACTTCAAGGCAATAAT 480
Db 460 GACTTCCAGCGCGCGGAGCCCTTACAAAGTTCTCGAGTTCGCTTACTTCAAGGCAATAAT 519
QY 481 CTCGCCAAAGCCATCGCGCTGTCTCACACCTTCTTACTGAGAGCATCTGTAGCAAAATG 540
Db 520 CTCGCCAAAGCCATCGCGCTGTCTCACACCTTCTTACTGAGAGCATCTGTAGCAAAATG 579
QY 541 ATGAAGAGAAACATGGCATATTTAAGAGCCTGCTGTGCGCGGAGGACTACATTAAAGAC 600
Db 580 ATGAAGAGAAACATGGCATATTTAAGAGCCTG-CTGTGCGCGGAGGACTACATTAAAGAC 638
QY 601 CTGGAACCAAGT--CATATGAAGCCCTGTTATCCCGAGAGTGGGGGACATACACGGTG 658
Db 639 CTGGAACCAAGTTCATATGAAGCCCTGTTATCCCGAGAGTGGGGGACATACACGGTG 698
QY 659 AGAAGTGGAGAACATCCATCACAGACATGGAGCTGGGCC--TTCCGACTTCTTCAAAGC 716
Db 699 AAAAGTGGAGAACATTCATCACAGACATGGAGCTGGGCCCTTCCGAAATTCCTTCAAAG 758
QY 717 CTTT 721
Db 759 NCTT 763

RESULT 7

ABL67662

ID ABL67662 standard; DNA; 2347 BP.

XX ABL67662;

XX 15-MAY-2002 (first entry)

XX Oesophagus cancer related gene sequence SEQ ID NO:5999.

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;

XX stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;

XX cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;

XX gene; ds.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

OS Homo sapiens.

PN WO200194629-A2.

XX 13-DEC-2001.

XX 30-MAY-2001; 2001WO-US10838.

XX 05-JUN-2000; 2000US-209473P.

XX 08-JUN-2000; 2000US-209531P.

XX 18-SEP-2000; 2000US-233113P.

XX 18-SEP-2000; 2000US-233617P.

XX 20-SEP-2000; 2000US-234009P.

XX 20-SEP-2000; 2000US-234034P.

XX 20-SEP-2000; 2000US-234052P.

XX 22-SEP-2000; 2000US-234509P.

XX 22-SEP-2000; 2000US-234567P.

XX 25-SEP-2000; 2000US-234923P.

XX 25-SEP-2000; 2000US-234924P.

XX 25-SEP-2000; 2000US-235077P.

XX 25-SEP-2000; 2000US-235082P.

XX 25-SEP-2000; 2000US-235134P.

XX 25-SEP-2000; 2000US-235280P.

XX 26-SEP-2000; 2000US-235637P.

XX 26-SEP-2000; 2000US-235638P.

XX 27-SEP-2000; 2000US-235711P.

XX 27-SEP-2000; 2000US-235720P.

XX 27-SEP-2000; 2000US-235840P.

XX 27-SEP-2000; 2000US-235853P.

XX 28-SEP-2000; 2000US-236028P.

XX 28-SEP-2000; 2000US-236032P.

XX 28-SEP-2000; 2000US-236033P.

XX 28-SEP-2000; 2000US-236034P.

XX 28-SEP-2000; 2000US-236109P.

XX 28-SEP-2000; 2000US-236111P.

XX 29-SEP-2000; 2000US-236842P.

XX 29-SEP-2000; 2000US-236891P.

XX 02-OCT-2000; 2000US-237172P.

XX 02-OCT-2000; 2000US-237173P.

XX 02-OCT-2000; 2000US-237278P.

XX 02-OCT-2000; 2000US-237294P.

XX 02-OCT-2000; 2000US-237295P.

XX 02-OCT-2000; 2000US-237316P.

XX 03-OCT-2000; 2000US-237425P.

XX 03-OCT-2000; 2000US-237598P.

XX 03-OCT-2000; 2000US-237604P.

XX 03-OCT-2000; 2000US-237606P.

XX 01-NOV-2000; 2000US-237608P.

XX 01-NOV-2000; 2000US-244867P.

XX 01-NOV-2000; 2000US-245084P.

XX (AVAL-) AVALON PHARM.

XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;

XX Soppet DR, Weaver Z;

XX WPI; 2002-188264/24.

XX Screening for anti-neoplastic agent involves exposing cells to a

XX chemical agent to be tested for anti-neoplastic activity, and

XX determining a change in expression of a gene of a signature gene set

XX Claim 1; SEQ ID 5999; 44pp; English.

XX The present invention describes a method (M1) for screening for an

XX anti-neoplastic agent. The method involves exposing cells to a chemical

XX agent to be tested for anti-neoplastic activity, determining a change in

XX expression of at least one gene (I) of a signature gene set, where (I)

XX comprises a sequence (S) selected from 8447 sequences (given in ABL61664

XX to ABL70110), or is at least 95% identical to (S), where a change in

XX expression is indicative of anti-neoplastic activity. (I) has cytostatic

XX activity and can be used in gene therapy. M1 can be used for screening

XX an anti-neoplastic agent, and can be used for producing a product which

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or properties of the agent. M1 can be used in the treatment of cancer such as colon, breast, stomach, lung, thyroid, esophageal, ovarian, kidney, prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms' tumour.

Sequence 2347 BP; 517 A; 689 C; 733 G; 408 T; 0 other;

Query Match 39.7%; Score 465.8; DB 24; Length 2347;

Best Local Similarity 64.2%; Pred. No. 2.3e-32;

Matches 752; Conservative 2; Mismatches 399; Indels 18; Gaps 3;

QY 14 GCCGGGGGGCGGGCGCTCTAGCGCTGTGTGCGTGGCTGGCGCTGGCGGGCGGC 73

DB 1 GCGCGGGCGGCGTGGCTGGCGTGGCGGCTGTGTGTTGCTGGGAGCGCG 60

QY 74 GCGCCCAATACGAACGCTACAGCTTCGAGCTTCCACGGGACGAGCTGATCCGCTCG 133

DB 61 GGGCGCAGTACGAGAGTACAGCTTCGGGGCTTCCGCCGAGGACCTGATCCCGCTGG 120

QY 134 AGTCGGCTACCGCGCGCGCTCGACAGTACAGCGCGGCGGAGCTGGCGGAGAGCTKG 193

DB 121 CCGCGCGGTACGGGCGCGCTCTGGAGCAGTACGAGGAGAGCTGGCGGAGCGCGC 180

QY 194 GCTACCTGGAGATCAGCTTCGGCTGCACCGCTTGTGTGCGGACAGCGAGGCTTCTGCC 253

DB 181 GCTACCTGGAGGCGGCGCTCGCGCTGCACCGCTCTCTGCGGACAGCGAGGCTTCTGCC 240

QY 254 ACCGCACTCGAGG-----CGCGCGGCGAGCGCGGCGCGCGCGCGCTGCCA 304

DB 241 ACGCCAACTCGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 300

QY 305 GCTATCC-----CGAGCTCGCGCTCTTCGGGGCGCTGTGCGCGCGCGCACTGCTCA 358

DB 301 ACCAGTGGGCGCTCGAGCTGGGCTCTTCGGCGCGCTCTTGAGCGAGCGCGCTGCTGC 360

QY 359 AGCGCTGCAAGCGGCGCTCGAGCTTCGCCAGCTCCAGCGCGCGCGCGGAGTCTGG 418

DB 361 GCGCTGCAAGCGGCGCGCTCGCGCGCTTCAGGTGCTTACCGCGCGCGCGAGTCTGC 420

QY 419 CGGACTTCCAGCGCGCGCGCGCTCAAGTTCCTCGAGTTCGCTTACTTCAAGGCAATA 478

DB 421 GTGACTTCCAGCGCGCGCTGCGCTACAGTACCTGACTACGCGCTGTTCAAGGCTAAC 480

QY 479 ATCTCCCAAGCCATCGCGCTGCTCAACCTTTCTACTGAAGCATCTGTAGCAAA 538

DB 481 GGCTGGAGAAGCGGTGGCGGCGCTACACCTTCTCCAGAGGAACCGAAGCAGCAGC 540

QY 539 TGATGAAGAGGACATGGCATATTATAA---GAGCGTCTGTCGCGGAGGACTACATTA 595

DB 541 TGACCGCGAGTATCTCACTACTATCAGGGGATGCTGGACGTGCGCGAGTCTCCTCA 600

QY 596 AAGACCTGGAAACCAAGTCAATGAAGCTTTCATCCGAGCAGTGGCGGATCAAG 655

DB 601 CGGACCTAGAGCGCGCGCTACGAGCGCTGTCTCTCGGCGTGTGAAGCTCTCAACA 660

QY 656 GTGAGAACTGGAGAACATCCATCAGAGATGAGCTGGCGCTTCCCGACTTCTTCAAG 715

DB 661 GCGGGATTTCGCGAGCAGCAGGAGCATCGAGCGGCGCTTGTGAGATACCTGGCAG 720

QY 716 CTTTTACAGTGTCTTCGCGAGGCTCGAGGGTTCCAGGGAGATCAAGACTTCAAGGATT 775

DB 721 TCTTTCGCGGTGCTGCGCGCTGTGAAGGGCGCATGAGCAGTGGACTTCAAGGACT 780

QY 776 TCTACCTTCCATAGCAGTCAATATGATGAAGTCTTGGAATGCAAAATACAGTGTGAG 835

DB 781 TCTACCGCGCCATAGCAGATCTCTTTCAGAGTTCCTGCAAGTGGAGTGTGAGG 840

QY 836 AGAACCTCACCCCGATTATAGGAGGCTATCCGGTTGAGAAATTTTGGCTACCATGATC 895

841 CCAATTGACCCCAATGTGGTGGCTACTTCGTGACAAAGTTCGTGGCCACCATGTACC 900
896 ATTACTTCAGATGGCTTATTAAGTTGAAGACCTGAAGATGAGCCCTCTGTGAG 955
901 ACTACTTCAGATGGCTTATTAAGTTGAAGATGAGCCCTCTGTGAG 960
956 TCAGCTATCTGCTCTTTGATCAGATGACAAAGTATGAGCAGAACTGGTGTATTACC 1015
961 CCAGCTACATGCTCTTCGACCCCAAGGACAGGCTATGAGCAGAACTGGTGTATTACC 1020
1016 AGTACACAGGACACTTGGGGCTCTCGGATGAGACTTCCAGCCGAGACTGAAGCAG 1075
1021 GGTTCACCGGCTCTGCTGGGGCTCTGAAGAGGAGGACTTCCAGCCGAGGAGGCA 1080
1076 TTCACTTCTTTAATGTGACCAACACCTCCAGAGGAGCTGTATGACTTTCTTAAGGAATA 1135
1081 TGCTCTACCAACAGCAGCCGCGAGCTGCGGAGCTGCTGGAGTTCCACCATGTACC 1140
1136 TAAATGATGATGATGAGGAGAGTGTGGA 1166
1141 TGCACTCAGATGATGAGATGGAGCTGGAGGA 1171

RESULT 8
AAH30585
ID AAH30585 standard; cDNA; 408 BP.
AC AAH30585;
XX
XX
27-JUL-2001 (first entry)
XX
DE Human colon cancer cell line Kml2L4-A cDNA library derived sequence #519.
XX
KW Human; diagnosis; colon cancer; cancer; malignant; chromosome mapping;
KW detection; colon cancer cell line Kml2L4-A; ss.
XX
OS Homo sapiens.
XX
PN WO200018916-A2.
XX
XX
06-APR-2000.
XX
XX 23-SEP-1999; 99WO-US22226.
XX
PR 28-SEP-1998; 98US-0102161.
PR 28-SEP-1998; 98US-0102180.
PR 29-SEP-1998; 98US-0102380.
PR 08-OCT-1998; 98US-0103815.
PR 27-OCT-1998; 98US-0105877.
XX
XX (CHIRON) CHIRON CORP.
XX (HYSE-) HYSEQ INC.
XX
XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
XX Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;
XX Lamson G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;
XX Leshkowitz D, Kita D, Garcia V, Jones LW, Stache-Crain B;
XX WPI; 2000-293155/25.
XX
XX Polynucleotide library comprising 1079 defined sequences, useful in
XX the form of an array to detect cancer or susceptibility to cancer -
XX
XX Claim 1; Page 333; 502pp; English.
XX
XX The present invention describes a library of polynucleotides comprising
XX 1079 nucleotide sequences (given in AAH30067 to AAH31145). Also described
XX are: (1) an isolated polynucleotide (I) having at least 90% identity to
XX one of the 1079 sequences; (2) a recombinant host cell containing (1);
XX (3) an isolated polypeptide (II) encoded by (1); (4) an antibody that
XX specifically binds to (II); (5) a vector comprising (1); and (6) a method
XX of detecting differentially expressed genes correlated with a cancerous
XX state of a mammalian cell comprising detecting a gene product encoded by

65 of the 1079 sequences given in the specification. The polynucleotides
are used to monitor patients having (or susceptible) to cancer to detect
potentially malignant events at a molecular level before they are
detectable at a gross morphological level. The polynucleotides are also
useful for monitoring the efficacy of various therapies and preventive
interventions. Polynucleotide probes based on the disclosed sequences
are useful for chromosome mapping and detection of transcription levels.
The 1079 polynucleotide sequences were derived from a human colon cancer
cell line Kml2L4-A cDNA library.

Sequence 408 BP; 117 A; 96 C; 98 G; 97 T; 0 other;

Query Match 33.8%; Score 406.4; DB 21; Length 408;
Best Local Similarity 99.8%; Pred. No. 1.5e-75;
Matches 407; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 497 CCCTGCTCACACCTTTCTACTGAAGCATCTGATGACGAATGATGAAGAGGAACATGG 556
Db 1 CCCTGCTCACACCTTTCTACTGAAGCATCTGATGACGAATGATGAAGAGGAACATGG 60
QY 557 CATATTATAAGAGCCTGCTGTCGCGAGGACTACATTAAAGACCTGGAAACCAAGTCAT 616
Db 61 CATATTATAAGAGCCTGCTGTCGCGAGGACTACATTAAAGACCTGGAAACCAAGTCAT 120
QY 617 ATGAAAGCCTGTTTCATCCGAGCAGTCCGGGCATACAAACGTCGAGACTGGAGAACATCCA 676
Db 121 ATGAAAGCCTGTTTCATCCGAGCAGTCCGGGCATACAAACGTCGAGACTGGAGAACATCCA 180
QY 677 TCACAGACATGAGTGGCCCTTCCGACTTCTTCAAAGCCTTTTACGAGTGTCTCGCAG 736
Db 181 TCACAGACATGAGTGGCCCTTCCGACTTCTTCAAAGCCTTTTACGAGTGTCTCGCAG 240
QY 737 CTGCGAGAGTTCAGGGAGATCAAGGACTTCAAGATTTTACCTTTCCATAGCAGATC 796
Db 241 CTGCGAGAGTTCAGGGAGATCAAGGACTTCAAGATTTTACCTTTCCATAGCAGATC 300
QY 797 ATTATGTAGAGTTCCTGCAATGCAAAATACAGTGTGAAGAGAACCTCACCCAGTTATAG 856
Db 301 ATTATGTAGAGTTCCTGCAATGCAAAATACAGTGTGAAGAGAACCTCACCCAGTTATAG 360
QY 857 GAGGCTATCCGTTGAGAAATTTGTGGCTACCATGTATCATCTATTC 904
Db 361 GAGGCTATCCGTTGAGAAATTTGTGGCTACCATGTATCATCTATTC 408

RESULT 9

ABZ11344
ID ABZ11344 standard; cDNA; 1309 BP.

XX ABZ11344;
XX
XX
XX 20-JAN-2003 (first entry)
XX
XX Human polynucleotide SEQ ID NO 226.

XX Human; genome mapping; gene therapy; food supplement; virus; fungus;
XX cell-proliferative disorder; neurodegenerative disease; bacterial;
XX Parkinson's disease; Alzheimer's disease; autoimmune disease;
XX multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
XX arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;
XX antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
XX haemostatic; vulnary; fungicide; antibacterial; virucide; protozoacide;
XX antiarthritic; gene; ss.
XX
XX Homo sapiens.

OS
XX
XX WO200270539-A2.

PN
XX
XX 12-SEP-2002.

PD
XX
XX 05-MAR-2002; 2002WO-US05095.

PF
XX
XX 05-MAR-2001; 2001US-0799451.

PR

XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
PI Wehrman T, Wang J, Wang D, Drmanac RT;
XX WPI: 2002-759812/82.
DR P-PSDB; ABP69127.
XX
PT New polynucleotides comprising sequences assembled from expressed
PT sequence tags (ESTs), useful for treating cell-proliferative,
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or
PT platelet or coagulation disorders
XX
PS Claim 1; SEQ ID NO 226; 1012pp + Sequence Listing; English.
XX
CC The invention relates to an isolated polynucleotide (I) comprising a
CC nucleotide sequence selected from any of 948 sequences
CC (AB211119-AB212066) or their mature protein coding portion, active domain
CC coding protein or complementary sequences. The polynucleotides are useful
CC for identifying expressed genes or for physical mapping of human genome.
CC The encoded polypeptides (ABP6902-ABP6949) are useful as molecular
CC weight markers, as a food supplement, for generating antibodies, in
CC medical imaging, screening and diagnostic assays and for treating
CC cell-proliferative disorders (cancer), neurodegenerative diseases
CC (Parkinson's or Alzheimer's disease), autoimmune diseases (multiple
CC sclerosis, diabetes, lupus) genetic disorders, myeloid or lymphoid
CC disorders, platelet or coagulation disorders, wound, burns, incision,
CC ulcers, liver or lung fibrosis, infections (bacterial, viral, fungal,
CC parasitic), arthritis, etc.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1309 BP; 287 A; 384 C; 365 G; 273 T; 0 other;

Query Match 13.0%; Score 156.4; DB 24; Length 1309;
Best Local Similarity 53.6%; Pred. No. 1.1e-24;
Matches 437; Conservative 0; Mismatches 346; Indels 33; Gaps 4;

QY 222 CCGTTGTCGGCGACGACGAGGCTTCTGCGACCGCACTGACGCGCGCGCGCGACCC 281
DB 256 CGCACACAGTGTGCGCGCGACTTCCGCTGGAGCTGGACCCCGACTGTCTCCCGACCC 315
QY 282 CGAGCCCGCGCGCTTCCGACGCTTCCGAGCTGCGCTTTCGGGGCGCTTCTGCG 341
DB 316 GCGCGAGCGCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 375
QY 342 CCGCGCGCGCTTCCGAGCGCTTCCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 401
DB 376 TCGGCTGCGCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 426
QY 402 CAGCGCGCGCTTCCG 461
DB 427 CAGCGAGAGATGAGCTGGAGTTCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 486
QY 462 TTACTTCAAGCAATATCTCCCGAAGCGCTTCCGCGCGCGCGCGCGCGCGCGCGCG 521
DB 487 CTACTTCAAGATCAAGTGGAGAGCTTGTGCTGCGAGCGCGCGCGCGCGCGCGCG 546
QY 522 GCATCTCTGATGACGAATATGAGAGGACATGGCATATTATAGAGCGCTGCTGGT- 580
DB 547 CAATCTTGAGCACATGGAATGACGAGACGACCTAGACTATTACCAACCAATGCTGAGT 606
QY 581 --CCGAGCTACATTAAGACCTGGAAACCAAGCTATATGAAGCGCTTCTATCCGAGC 638
DB 607 GAAGAGGCGCGACTTCAAGGATCTTGGAGCTCAACCCCATATGCAAGAAATTCGACTGG 666
QY 639 AGTCGGGCGCATACAAACGCTGAGAACTGGAGAAATCCATCCATCAGACATGAGCTGCGCT 698
DB 667 AGTCGAGCTTCTCAGAGGACGAGCCACAGGAGAGCTGTGCCCCACCTAGAGCGCGCT 726

699 TCCGACTTCTCAAAGCCCTTTACGAGTGTCTCGCAGCGCTCGAGGGTTCAGGAGAT 758
727 GCAAGATATCTTGTGCGCTATGAGAGTGCCTGCGCTTGCAGAGGCGCTTATGACTA 786
759 CAGGACTTCAA-----GGATTCTACCTTTCCATAGCAGATCATTA 800
787 CGATGGCTACAACTACCTTGAGTACAAACGCTGACCTCTCCAGGCCATCACAGATCATTA 846
801 TGTAGAAATTTCTGGAATGCAAAATACAGTGTGAAGAGAACCT--CACCCAGTTATAGG 857
847 CATCCAGGTCTCTCACTGTAAAGCAGAACTGTGTACGAGGCTTGTCCACCCAGTCG 906
858 AGCGTATCCGTTGAGAAATTTGTGCTACCAATGTATCATTAATTCAGTTGCTTATTA 917
907 AGAAGAGCGCTTTGAAGACTTCTCCCATCGCAATTAATATCTCAGTTTGGCTTACTA 966
918 TAAGTTGAACGACCTGAAGAATGCAGCCCGCTGTGCGAGTCAGCTATCTCTTTGATCA 977
967 TAACATGGGAATATACAGGCTTTGAATGTGCCAAGACCTATCTCTCTTCTTCTCC 1026
978 GAATGACAGGTTCATCCAGCAGAACCTGTGTATTA 1013
1027 CAATGACGAGGTGATGAACCAAAATTTGGCTTATTA 1062

RESULT 10
ABZ11345
ID ABZ11345 standard; cDNA; 2152 BP.
XX AC ABZ11345;
XX AC
XX AC
DT 20-JAN-2003 (first entry)
XX
DE Human polynucleotide SEQ ID NO 227.
XX
XX Human; genome mapping; gene therapy; food supplement; virus; fungus;
XX cell-proliferative disorder; neurodegenerative disease; bacterial;
XX Parkinson's disease; Alzheimer's disease; autoimmune disease;
XX multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
XX arthritis; cytostatic; immunomodulator; neutropenic; neuroprotective;
XX antiparkinsonian; antididiabetic; immunosuppressive; dermatological;
XX haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
XX antiarthritic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200270539-A2.
XX
PD 12-SEP-2002.
XX
PF 05-MAR-2002; 2002WO-US05095.
XX
XX 05-MAR-2001; 2001US-0799451.
PR
XX (HYSE-) HYSEQ INC.
XX
PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
PI Wehrman T, Wang J, Wang D, Drmanac RT;
XX
XX WPI: 2002-759812/82.
DR P-PSDB; ABP69127.
XX
PT New polynucleotides comprising sequences assembled from expressed
PT sequence tags (ESTs), useful for treating cell-proliferative,
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or
PT platelet or coagulation disorders
XX
PS Claim 1; SEQ ID NO 227; 1012pp + Sequence Listing; English.
XX
CC The invention relates to an isolated polynucleotide (I) comprising a
CC nucleotide sequence selected from any of 948 sequences
CC (AB211119-AB212066) or their mature protein coding portion, active domain

CC coding protein or complementary sequences. The polynucleotides are useful
CC for identifying expressed genes or for physical mapping of human genome.
CC The encoded polypeptides (ABP9902-ABP9949) are useful as molecular
CC weight markers, as a food supplement, for generating antibodies, in
CC medical imaging, screening and diagnostic assays and for treating
CC cell-proliferative disorders (cancer), neurodegenerative diseases
CC (Parkinson's or Alzheimer's disease), autoimmune diseases (multiple
CC sclerosis, diabetes, lupus) genetic disorders, myeloid or lymphoid
CC disorders, platelet or coagulation disorders, wound, burns, incision,
CC ulcers, liver or lung fibrosis, infections (bacterial, viral, fungal,
CC parasitic), arthritis, etc.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2152 BP; 494 A; 636 C; 602 G; 420 T; 0 other;
Query Match 13.0%; Score 156.4; DB 24; Length 2152;
Best Local Similarity 53.6%; Pred. No. 1.2e-24; Indels 33; Gaps 4;
Matches 437; Conservative 0; Mismatches 346;
QY 222 CGCTTGTCTGCGCAGCGAGCGCTTCTGCCACGCAACTGCGCGCGCGCGAGCC 281
Db CGCACCAGTGTGCGCGCGACTTCCCGTGGAGCTGAGCCCGCGACTGTCGCCAGCC 315
QY 282 CGAGCCGCGCGCGCTCGCAGTATCCGAGTGTGCGCTTTCGGGGCGCTGCTGG 341
Db GCGCCAGGCTCGGGCGCGCGCTCGGAGCTTCTGCGGGCGCTTCTGCG 375
QY 342 CGCGCGCACTGCTCAAGCGGTGCAAGCGGGCTGCGAGCTTCCGCCAGTCCCGAGCC 401
Db TCGCGTGTCTGCGCGCGCTGCTCGGCGCGCGCGCGCGCTC-----GCT 426
QY 402 CAGCGCGAGTGTGCGGACTTCCAGCGCGCGCGAGCTTCAAGTTCCTGAGTTCGC 461
Db CAGCGAGAGATGAGTGTGAGTTCGCGAGCGAGCGCGCTTCACTACCTGCGAGTGC 486
QY 462 TTACTTCAAGGCAATATCTCCCAAGGCACTGCGCGCTGCTCACACCTTCTTACTGAA 521
Db CTACTTCAAGATCAACAGTGTGGAGAAAGCTGTGCTGCAGACACACCTTCTTCGTGG 546
QY 522 GCATCTGTATGCAATGATGAGGAGCAATGTCATATTATGAGCGCTGCTGCTGTG- 580
Db CAACTCTGAGCATGGAATGCAAGAACCTAGACTATTACCAACCATGCTGCGAGT 606
QY 581 --CCGAGGACTATTAAAGACTGGAACCAAGTCATATGAAAGCCTGTTTCATCCGAGC 638
Db GAAGGAGCGGACTTCAAGATCTTGAAGTCAACCCCATATGCAAGAAATTGCGAGTGG 666
QY 639 AGTGGGCTATCAAGCTGAGTGAATGAGBACATCCATCAGACATGAGCTGGCCCT 698
Db AGTGGACTCTACTCAGAGGAACGCCACAGGAAGCTGTGCCACCTAGAGCGCGGCT 726
QY 699 TCCGACTTCTTCAAGCGCTTTTACAGTGTCTGCGAGCTGCGAGGCTTCAGGGAGAT 758
Db GCAAGAATCTTGTGGCTATGAGGAGTGGCGCTGCTGCGAAGGCGCTATGACTA 786
QY 759 CAAGACTTCAA-----GGATTCTACCTTCCATAGAGTCAATTA 800
Db CGATGGCTCAACTACTCTTGAATCAACCGCTGACCTTTCAGGCGCATCAGATCATTA 846
QY 801 TGTAGAAGTCTTGAATGCAAAATACAGTGTGAAGAGAACCT---CCCCCAGTTATAGG 857
Db CATCCAGTCTCACTGTAAGCAGACTGTGTACAGGAGCTTGTCTCCCAACCAAGTCG 906
QY 858 AGGCTATCGGTTGAGAAATTTGGCTACCATGTATCATTTACTTGAGTTGCTTATTA 917
Db AGAAGAGCCCTTGAAGACTTCTCCCATCGCAITTAATTAATTCATGCGAGTTGCTACTA 966
QY 918 TAAGTTGAAGCAGCTGGAAGATGCGAGCCCGCTGTGCGAGTCACTATCTGCTTCTTATCA 977
Db TAACATTGGGAATTATACACAGCGCTGTTGAATGTGCCAAGACCTATCTTCTTCTCCC 1026

QY 978 GAATGCAAGGTCATGCGAGCAGACACCTGGTGTATTA 1013
Db 1027 CATGACGAGTGAAGAACCAAAATTGGCTTATTA 1062
RESULT 11
AAS17573
ID AAS17573 standard; cDNA; 2211 BP.
XX AAS17573;
AC AAS17573;
XX
DT 26-FEB-2002 (first entry)
XX
DE DNA encoding novel secreted protein #2.
XX
KW Secreted protein; cytostatic; immunosuppressive; vulnery; vaccine;
KW antinflammatory; neuroprotective; nephrotropic; cardiovascular;
KW human; cancer; autoimmune disease; wound healing disorder; infection;
KW haematopoietic disorder; inflammatory disorder; infertility;
KW neurological disease; psychiatric disease; cardiovascular disease;
KW respiratory disease; renal; gastrointestinal; ss.
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT 1.2211
FT CDS /*tag= a
FT /product= "Human secreted protein"
XX WO200179454-A1.
XX
XX 25-OCT-2001.
XX
XX 11-APR-2001; 2001WO-US11797.
XX 13-APR-2000; 2000US-196603P.
XX 24-APR-2000; 2000US-199417P.
XX (SMIK) SMITHKLINE BEECHAM CORP.
PA (SMIK) SMITHKLINE BEECHAM PLC.
XX
XX Agarwal P, Murdoch PR, Rizvi SK, Smith RP, Xiang Z;
XX WPI; 2002-061975/08.
DR P-PSDB; AAU09861.
XX
XX New secreted proteins or polypeptides, useful for treating e.g. cancer,
XX autoimmune diseases, wound healing disorder, infections, haematopoietic
XX disorders, inflammatory disorders, infertility, cancer -
XX Claim 2; Page 35-36; 92pp; English.
XX
XX The invention relates to an isolated novel secreted polypeptide (I) and
XX polynucleotide (II). (I) and (II) are useful for treating cancer,
XX autoimmune diseases, wound healing disorder, infections, haematopoietic
XX disorders, inflammatory disorders, infertility, neurological and
XX psychiatric diseases, cardiovascular diseases, respiratory diseases,
XX renal diseases, or gastrointestinal diseases. These may also be used to
XX treat diseases, abnormalities and disorders caused by abnormal
XX expression, production, function and/or metabolism of the genes, as
XX vaccines for inducing immunological response in a mammal, and in
XX screening methods for detecting the effect of added compounds on the
XX production of mRNA and polypeptide in cells. The polypeptides can be used
XX as immunogens to produce antibodies immunospecific for the polypeptides,
XX and to identify membrane-bound or soluble receptors. The polynucleotides
XX may be used as diagnostic reagents, in chromosome localisation studies,
XX and in tissue expression studies. The present sequence represents the
XX coding sequence of novel human secreted protein #2.
XX
SQ Sequence 2211 BP; 512 A; 649 C; 612 G; 438 T; 0 other;
Query Match 13.0%; Score 156.4; DB 24; Length 2211;
Best Local Similarity 53.6%; Pred. No. 1.2e-24;

Db 771 CGATGGCTCAACTACTTGTAGTACAAAGCTGACCTTCCAGGCGCATCAGATCATTA 830
 Qy 801 TGTAGAAGTTCTGGATGCAAAATACAGTGTGAAGAGAACCT---CACCCAGTTATAGG 857
 Db 831 CATCCAGGCTCTCAACTGTGAAGCAGAACTGTGTACGGAGCTGTCTTCCACCCCAAGTCG 890
 Qy 858 AGGCTATCCCGTGTGAGAAATTTGTGCTACCATGTATCATTTACTTGCAGTTTGCCTATTA 917
 Db 891 AGAGAGCCCTTTGAAGACCTTCTCCCATCGCATTAATATCTCGAGTTTGCCTACTA 950
 Qy 918 TAGATTGAACGACCTGAAGAAATGACGCCCCCTGTGAGTCAGCTATCTGCTTTGATCA 977
 Db 951 TAAATATGGGAATATACACAGCTGTGAATGTGCAAGACCTATCTTCTTCTTCTCC 1010
 Qy 978 GAATGACAAAGTCATGACGAGAACCTGGTGTATTA 1013
 Db 1011 CAATGACGAGTGTATGAACCAAAATTTGGCCTATTA 1046

 RESULT 14
 AAF93800
 AAF93800 standard; cDNA; 2563 BP.
 AC AAF93800;
 XX
 DT 23-MAY-2001 (first entry)
 XX
 DE Human cDNA encoding a membrane or secretory protein clone P85C0109.
 XX
 KW Human; secretory protein; membrane protein; vaccine; gene therapy;
 KW rheumatoid arthritis; diabetes; ss.
 XX
 OS Homo sapiens.
 XX
 PN EPI067182-A2.
 XX
 PD 10-JAN-2001.
 XX
 PF 07-JUL-2000; 2000EP-0114090.
 XX
 PR 08-JUL-1999; 99JP-0194179.
 PR 11-JAN-2000; 2000JP-0118775.
 PR 02-MAY-2000; 2000JP-0183766.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T. Isogai T. Nishikawa T. Kawai Y. Sugiyama T. Hayashi K;
 XX
 DR WPI: 2001-093989/11.
 DR P-PSDB; AAB88373.
 XX
 PT Nucleic acids encoding secretory proteins/membrane proteins, useful in
 PT gene therapy or as candidate target molecules in drug development -
 XX
 PS Claim 1; SEQ ID 113; 609pp + CD ROM; English.
 XX
 CC This invention relates to nucleic acid sequences AAF93744 - AAF93916
 CC which encode human secretory or membrane proteins represented by
 CC AAB88317 - AAB88419. Included in the invention are primers
 CC AAF93917 - AAF94295 and AAF62232 - AAF62235 which are used to isolate the
 CC cDNA sequences of the invention. The invention also includes methods for
 CC the production of antibodies directed against the proteins, and cDNA
 CC sequences, which can be used in vaccines. The polynucleotide sequences
 CC can be used in gene therapy. The polynucleotide sequences and the
 CC proteins they encode may be used in the prevention, treatment and
 CC diagnosis of diseases associated with inappropriate secretory
 CC protein/membrane protein expression. The nucleic acids and complementary
 CC sequences may also be used as DNA probes in diagnostic assays
 CC (e.g. polymerase chain reactions (PCR)) to detect and quantitate the
 CC presence of similar nucleic acid sequences in samples. They may also be
 CC used to study the expression and function of secretory proteins/membrane
 CC polypeptides and their role in metabolism. The polypeptides may be used
 CC as antigens in the production of antibodies against them and in assays to

CC identify modulators (agonists and antagonists) of expression and
 CC activity. The antibodies and antagonists may also be used as therapeutic
 CC agents to down regulate expression and activity. The antibodies may also
 CC be used as diagnostic agents for detecting the presence of the
 CC polypeptides in samples (e.g. by enzyme linked immunosorbant assay
 CC (ELISA)). Examples of diseases which may be treated include rheumatoid
 CC arthritis and diabetes.
 XX
 SQ Sequence 2563 BP; 589 A; 753 C; 713 G; 508 T; 0 other;

 Query Match 13.0%; Score 156.4; DB 22; Length 2563;
 Best Local Similarity 53.6%; Pred. No. 1.2e-24;
 Matches 437; Conservative 0; Mismatches 346; Indels 33; Gaps 4;

 Qy 222 CGCGTTGTCTGCGCAGCAGCGAGGCGCTTCTGCACACGCAACTGCGAGCGCGCGCGAGCC 281
 Db 255 CGGACCCAGTGTGCGCGGACTTCCCGTGGAGCTGGACCCCGGACCTGGTCCCGCCAGCC 314
 Qy 282 CGAGCCCGCGCGCGGCTGCGCAGCTATCCGAGTGCGGCTCTTTCGGGGCGCTGTGCG 341
 Db 315 GGCCACAGGCTCGGGCGCGCGCGCTGCGGAGCTGAGCTTCTTCGGGGCGCTTCTGCG 374
 Qy 342 CGCGCGCAGCTGCTCAAGCGCTGCAAGCAGGCGCTGCCAGCCTTCCGCGAGTCCCGAGCC 401
 Db 375 TCGGCGTGTGCTGCGCGGCTGCTCGGCGCGCGCGCGCCCTC-----GCT 425
 Qy 402 CAGCCCGCAGGTGCTGGCGGACTTCCAGCGCGCGGAGCCCTACAAGTTCCTGCAATTCGC 461
 Db 426 CAGCGAAGAGATGGAGCTGGAGTTCCGCAAGCGGAGCGCCCTACAACTACCTGCGAGTGC 485
 Qy 462 TTACTTCAAGGCAAAATATCTCCCAAGCCATCGCGCTGCTCACACCTTTCTACTGAA 521
 Db 486 CTACITCAGATCAACAAGTTGGAGAAAGCTGTTCTGCASACACACCTTCTTCGTGGG 545
 Qy 522 GCATCTCTGATGAGAAATGATGAAGAGAAACATGCGCATATTATTAAGAGCCTGCTGTG 580
 Db 546 CAATCCTGAGCACATGGAAATGCAGCAGAACCTAGACTATTACCAACCATGTCTGGAGT 605
 Qy 581 --CCGAGGACTACATTAAAGACCTGGAACCAAGTCATATGAAAGCCTGTTCACTCCAGC 638
 Db 606 GAAGGAGCGGACTTCAAGGATCTTGAGACTCAACCCCATATGCAAGATTTGAGCTGGG 665
 Qy 639 AGTGGGGGATACAAAGCGGTGAGAACTGGAGAACTCATCATCACACATGAGAGTGGCCCT 698
 Db 666 AGTGGGACTCTACTCAGAGGAAACAGCCACAGAAAGCTGTGCCCCCACCCTAGAGGCGGCGCT 725
 Qy 699 TCCGACTTCTTCAAGGCTTTTACGAGTGTCTCGAGCCTGCGAGGTTCCAGGGAGAT 758
 Db 726 GCAAGNATACITTTGGGCTATGAGGAGTGGCGTGCCTCTGCGAAGGGCCCTATGACTA 785
 Qy 759 CAAGGACTTCAA-----GGATTTCTACCTTTCCATAGCAGATCATTA 800
 Db 786 CGATGGCTACAACTACCTTGAGTACAGCGCTGACCTCTTCAGGCCATCACAGATCATTA 845
 Qy 801 TGTAGAAGTTCTGGAAATGCAAAATACAGTGTGAAGAGAACCT---CACCCCGAGTTATAGG 857
 Db 846 CATCCAGTCTCAACTGTAGCAGAACTGTGTACGGAGCTTGTCTCCACCCAGTCG 905
 Qy 858 AGCTATCCGGTTGAGAAATTTGGTGCTACCATGTATCATTTACTTGCAGTTTGCCTATTA 917
 Db 906 AGAGAAGCCCTTTGAAGACTTCTCTCCCATCATTAATAATTATCTGAGTTTGCCTACTA 965
 Qy 918 TAAGTTGAACGACCTGGAAGAAATGACGCGCCCTGTGCACTGAGCTATCTGCTTTGATCA 977
 Db 966 TAACTTGGAAATTTATACAGGCTGTTGAATGTGCGCAAGACCTATCTTCTTCTTCCC 1025
 Qy 978 GAATGACAAAGTGTGAGCAGAGAACCTGGTGTATTA 1013
 Db 1026 CAATGACGAGTGTATGAACCAAAATTTGGCCTATTA 1061

 RESULT 15
 AAF14589

AAH14589 standard; cDNA; 2585 BP.
AAH14589;
26-JUN-2001 (first entry)
Human cDNA sequence SEQ ID NO:12193.
Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
Homo sapiens.
EP1074617-A2.
07-FEB-2001.
28-JUL-2000; 2000EP-0116126.
29-JUL-1999; 99JP-0248036.
27-AUG-1999; 99JP-0300253.
11-JAN-2000; 2000JP-0118776.
02-MAY-2000; 2000JP-0183767.
09-JUN-2000; 2000JP-0241899.
(HELI-) HELIX RES INST.
Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
WPI; 2001-318749/34.
Primer sets for synthesizing polynucleotides, particularly the 5602
full-length cDNAs defined in the specification, and for the detection
and/or diagnosis of the abnormality of the proteins encoded by the
full-length cDNAs -
Claim 8; SEQ ID 12193; 2537pp + CD ROM; English.
The present invention describes primer sets for synthesizing 5602
full-length cDNAs defined in the specification. Where a primer set
comprises: (a) an oligo-dT primer and an oligonucleotide complementary
to the complementary strand of a polynucleotide which comprises one of
the 5602 nucleotide sequences defined in the specification, where the
oligonucleotide comprises at least 15 nucleotides; or (b) a combination
of an oligonucleotide comprising a sequence complementary to the
complementary strand of a polynucleotide which comprises a 5'-end
sequence and an oligonucleotide comprising a sequence complementary to a
polynucleotide which comprises a 3'-end sequence, where the
oligonucleotide comprises at least 15 nucleotides and the combination of
the 5'-end sequence/3'-end sequence is selected from those defined in
the specification. The primer sets can be used in antisense therapy and
in gene therapy. The primers are useful for synthesizing polynucleotides,
particularly full-length cDNAs. The primers are also useful for the
detection and/or diagnosis of the abnormality of the proteins encoded by
the full-length cDNAs. The primers allow obtaining of the full-length
cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
represent oligonucleotides, all of which are used in the exemplification
of the present invention.
Sequence 2585 BP; 590 A; 759 C; 723 G; 513 T; 0 other;

Search completed: January 30, 2004, 08:32:16
Job time : 317.764 secs

Query Match 13.0%; Score 156.4; DB 22; Length 2585;
Best Local Similarity 53.6%; Pred. No. 1.2e-24;
Matches 437; Conservative 0; Mismatches 346; Indels 33; Gaps 4;
QY 222 CCGCTTCTCGGACAGCAGGCGCTTCTCCACCGCAACTGACGAGCGCGCGCAGCC 281
DB 277 CCGCACCAGTGTGCGCGCCGACCTCCCGTGGAGCTGGACCCGACTGTGTCGCCAGCCC 336
QY 282 CGAGCGCGCGCGCGCTCCGAGCTATCCGAGCTCGGCTCTTGGGGGCGCTGTGCG 341

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM nucleic - nucleic search, using sw model

Run on: January 30, 2004, 05:50:18 ; Search time 1006.41 Seconds
(without alignments)
10382.969 Million cell updates/sec

Title: US-09-729-674-1
Perfect score: 3871
Sequence: 1 tttctctctctctctctt.....ataaaaaaaaaaaaaaaaaa 3871

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*
25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2004.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3870.2	100.0	3871	20	Human secreted pro
2	3870.2	100.0	3871	22	Human cDNA encodin
3	3870.2	100.0	3871	24	Human polynucleoti
4	2051.2	53.0	2549	21	Human prostate can
5	1955.4	50.5	1961	22	Human cDNA sequenc
6	1474.4	38.1	1886	21	Human colon cancer
7	1062	27.4	2499	23	Human prostate exp
8	1062	27.4	2499	23	Human prostate exp

9	695.6	18.0	818	22	Human cDNA clone (
10	587.4	15.2	660	24	Human colon cancer
11	466.4	12.0	618	21	Human colon cancer
12	465.8	12.0	2347	24	Oesophagus cancer
13	406.4	10.5	408	21	AAH30585
14	392.4	10.1	394	24	AAH50786
15	371.6	9.6	401	22	AAH50786
16	369.4	9.5	400	16	AAH21281
17	355.6	9.2	407	22	AAH50808
18	340	8.8	351	22	AAH11840
19	337.2	8.7	373	21	AAH43677
20	331.2	8.6	595	23	Human prostate exp
21	267.6	6.9	350	23	Human prostate exp
22	263.4	6.8	274	24	ABL65995
23	258.6	6.7	279	23	AAH58007
24	241.2	6.2	43545	24	AAH85018
25	239.2	6.2	6670	22	AAH02846
26	239.2	6.2	7713	22	AAH02845
27	238.4	6.2	8658	22	AAH69231
28	238	6.1	295	22	AAH82995
29	235.4	6.1	32351	21	AAH21307
30	235.4	6.1	32351	21	AAH35185
31	235.4	6.1	40298	21	AAH21311
32	235.4	6.1	40298	21	AAH35189
33	234.6	6.1	4275	22	ABA3408
34	234.6	6.1	4883	22	AAH9410
35	234.6	6.1	8894	23	AAH42719
36	234.6	6.1	16997	22	AAH80033
37	234.6	6.1	16997	22	AAH86146
38	234.2	6.1	9163	24	ABL45808
39	234.2	6.1	97835	24	ABL45808
40	233.8	6.0	9620	22	AAH66207
41	233.6	6.0	43950	22	AAH36022
42	233.4	6.0	6138	22	AAH57456
43	233.2	6.0	763	22	AAH06411
44	232.6	6.0	2263	22	AAH14438
45	232.6	6.0	5979	22	AAH78985

ALIGNMENTS

RESULT 1
AAH60801
ID AAX60801 standard; DNA; 3871 BP.
XX
AC AAX60801;
XX
DT 09-AUG-1999 (first entry)
XX
DE Human secreted protein encoding DNA (clone bd306-7).
XX
KW Secreted protein; kidney; lung; brain; blood; testis; bone marrow;
KW nutritional activity; cytokine; cell proliferation; immune stimulation;
KW hematopoiesis regulation; tissue growth; thrombolytic; gene therapy;
KW anti-inflammatory; tumour invasion; ss.
OS Homo sapiens.
XX
PN WO9926961-A1.
XX
PD -03-JUN-1999.
XX
PF 24-NOV-1998; 98WO-US25149.
XX
PR 23-NOV-1998; 98US-0197886.
XX
PR 26-NOV-1997; 97US-0066804.
XX
FA (GENY) GENETICS INST INC.
XX
PI Agostino MJ, Clark HF, Collins-Racie LA, Evans C;
PI Fechtel K, Jacobs K, Lavallie ER, McCoy JM, Merberg D;
PI Steininger RJ, Treacy M, Wong GG;

1681 DB ||||| CCAGGCTGCTCGAACTCTTGACCTTCAGATGATCCATCTGCCCTGGCCCTCCACAGTGC 1740
1741 QY ||||| TGGGATTTACAGCGGTGAGCCACCATGCGCGGCTCTTTCTCACCTTTACACCTGCTCTCT 1800
1741 DB ||||| TGGGATTTACAGCGGTGAGCCACCATGCGCGGCTCTTTCTCACCTTTACACCTGCTCTCT 1800
1801 QY ||||| TATCCTCACATCTGTTTTACACCTTCATPCCTGCTCTTCTTCCTCATCTGTTACACCTGCTCT 1860
1801 DB ||||| TATCCTCACATCTGTTTTACACCTTCATPCCTGCTCTTCTTCCTCATCTGTTACACCTGCTCT 1860
1861 QY ||||| CCCATGTTTCATAGCTGCTCTTCTTACCATTTTGGTTTGAAGGAGCTCTCTCGGCTTG 1920
1861 DB ||||| CCCATGTTTCATAGCTGCTCTTCTTACCATTTTGGTTTGAAGGAGCTCTCTCGGCTTG 1920
1921 QY ||||| TTTTTTGTGTTTTCCAGAAATCAGTATATTTTTTAAATAGAAATTAATCTCTAGAA 1980
1921 DB ||||| TTTTTTGTGTTTTCCAGAAATCAGTATATTTTTTAAATAGAAATTAATCTCTAGAA 1980
1981 QY ||||| GATGATAATTTGAAACCTCTTGGCTTATTTGCTTTTCCAGATTTTAGTCTCTCTTC 2040
1981 DB ||||| GATGATAATTTGAAACCTCTTGGCTTATTTGCTTTTCCAGATTTTAGTCTCTCTTC 2040
2041 QY ||||| TCCCATCCGGAAGATGTTGGAAGACATAGGCTAAATTTCTCCAGCTCACAAATGCTC 2100
2041 DB ||||| TCCCATCCGGAAGATGTTGGAAGACATAGGCTAAATTTCTCCAGCTCACAAATGCTC 2100
2101 QY ||||| TTTCACTTGTCTGACTTGTACCAATTTCTAGCACCCACTGAAATTAAGTTGAGTAGAGAG 2160
2101 DB ||||| TTTCACTTGTCTGACTTGTACCAATTTCTAGCACCCACTGAAATTAAGTTGAGTAGAGAG 2160
2161 QY ||||| TGTAGAGTCAGAAATGTTGGCTTTTGGCCCTTGTGCTATCTCCAAATTAAGTTGAGTAGAG 2220
2161 DB ||||| TGTAGAGTCAGAAATGTTGGCTTTTGGCCCTTGTGCTATCTCCAAATTAAGTTGAGTAGAG 2220
2221 QY ||||| CCGATCCCATTTGAGGACAACTCTAGTTTAAAGTCTCCGAGTTGGAAGGAGGAGGAGC 2280
2221 DB ||||| CCGATCCCATTTGAGGACAACTCTAGTTTAAAGTCTCCGAGTTGGAAGGAGGAGGAGC 2280
2281 QY ||||| CAGAGCTGCTAGTTTCACTTCACTTCTTCACTTCACTTCACTTCACTTCACTTCACTT 2340
2281 DB ||||| CAGAGCTGCTAGTTTCACTTCACTTCTTCACTTCACTTCACTTCACTTCACTTCACTT 2340
2341 QY ||||| TAGGCAATGAGCTGGAATCTAGATACCTTACAGAAATTAAGTTGAGTAGAGTAGAGTAG 2400
2341 DB ||||| TAGGCAATGAGCTGGAATCTAGATACCTTACAGAAATTAAGTTGAGTAGAGTAGAGTAG 2400
2401 QY ||||| CATGGAGCTTACATTTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2460
2401 DB ||||| CATGGAGCTTACATTTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2460
2461 QY ||||| AAGGTATCATGTAGTATGATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2520
2461 DB ||||| AAGGTATCATGTAGTATGATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2520
2521 QY ||||| TCTGAGATGAGGCTGAGTTTAAATGAGGCTTCACTGGAATGTGACCTTGGAGGAGGAG 2580
2521 DB ||||| TCTGAGATGAGGCTGAGTTTAAATGAGGCTTCACTGGAATGTGACCTTGGAGGAGGAG 2580
2581 QY ||||| GAGCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2640
2581 DB ||||| GAGCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2640
2641 QY ||||| CCCCCAAGTGAGACTTGTCTCTACGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2700
2641 DB ||||| CCCCCAAGTGAGACTTGTCTCTACGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2700
2701 QY ||||| TGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2760
2701 DB ||||| TGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2760
2761 QY ||||| CTTCTGCTAAGTGAGATGGAAACCACTGGAGGTTTGAACAGAGAGGAGGAGGAGGAGGAG 2820

2761 DB ||||| CTTCTGCTAAGTGAGATGGAAACCACTGGAGGTTTGAACAGAGAGGAGGAGGAGGAGGAG 2820
2821 QY ||||| TTATATTTTGAAGGCTCATTTCTAGCTGCAATATTTGTAATAAATCTTTAGTGGACAAGGCG 2880
2821 DB ||||| TTATATTTTGAAGGCTCATTTCTAGCTGCAATATTTGTAATAAATCTTTAGTGGACAAGGCG 2880
2881 QY ||||| AGAAGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2940
2881 DB ||||| AGAAGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2940
2941 QY ||||| CACACCAACAGCAGTGGTCAATATCTAGATTTTATTTTGAAGGAGGAGGAGGAGGAGGAGG 3000
2941 DB ||||| CACACCAACAGCAGTGGTCAATATCTAGATTTTATTTTGAAGGAGGAGGAGGAGGAGGAGG 3000
3001 QY ||||| TGAGAGTTTGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3060
3001 DB ||||| TGAGAGTTTGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3060
3061 QY ||||| GCTCAATAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3120
3061 DB ||||| GCTCAATAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3120
3121 QY ||||| AAGTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3180
3121 DB ||||| AAGTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3180
3181 QY ||||| AATCAAGTGGAGAGTCTCTCAGGCTCAGGCTCGAGCAGAGGCTCGAGCAGAGGAGGAGGAG 3240
3181 DB ||||| AATCAAGTGGAGAGTCTCTCAGGCTCAGGCTCGAGCAGAGGCTCGAGCAGAGGAGGAGGAG 3240
3241 QY ||||| TGCACTTGGTTTATTTGTTGGGGTCTCTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3300
3241 DB ||||| TGCACTTGGTTTATTTGTTGGGGTCTCTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3300
3301 QY ||||| ATTTATTTGGCTGAGAGGAGTCTCTGGAGTGGTGGTACATTTGAGGAGGAGGAGGAGGAGG 3360
3301 DB ||||| ATTTATTTGGCTGAGAGGAGTCTCTGGAGTGGTGGTACATTTGAGGAGGAGGAGGAGGAGG 3360
3361 QY ||||| GTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3420
3361 DB ||||| GTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3420
3421 QY ||||| TGTGTTGGGGTGGGGTCTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3480
3421 DB ||||| TGTGTTGGGGTGGGGTCTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3480
3481 QY ||||| GGGCCCCCTACTGAACCAATCAGCATGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3540
3481 DB ||||| GGGCCCCCTACTGAACCAATCAGCATGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3540
3541 QY ||||| CACTGAGATTTGCTCTGAGTATTTACTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3600
3541 DB ||||| CACTGAGATTTGCTCTGAGTATTTACTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3600
3601 QY ||||| TTGGGCTCTCTGGGAAATGGGAAATCAGCCAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3660
3601 DB ||||| TTGGGCTCTCTGGGAAATGGGAAATCAGCCAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3660
3661 QY ||||| AGAGAAAACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3720
3661 DB ||||| AGAGAAAACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3720
3721 QY ||||| CTTCTCCAAATAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3780
3721 DB ||||| CTTCTCCAAATAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3780
3781 QY ||||| AAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3840
3781 DB ||||| AAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3840
3841 QY ||||| TAGAGTGTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3871
3841 DB ||||| TAGAGTGTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3871

RESULT 2

AAS59207

ID AAS59207 standard; cDNA; 3871 BP.

AC AAS59207;

XX

DT 16-

1. 2. 3. 4. 5.

DE Human cDNA encoding a secreted protein bd306_7.

Human; secreted protein; ss; antiinflammatory; immunosuppressive; cytotoxic; neuroprotective; antiarthritic; antimicrobial; vulnery; cytostatic; antidiabetic; virucide; antifertility; anticonvulsant; vasotropic; antiparkinsonian; immunostimulant; dermatological; antirheumatic; antitumor; antiulcer; osteopathic; tranquiliser; cerebroprotective; cytokine; cell proliferation; cell differentiation; immune deficiency; severe combined immunodeficiency; SCID; tumour; autoimmune disorder; multiple sclerosis; rheumatoid arthritis; graft-versus-host disease; myeloid deficiency; wound healing; ulcer; periodontal disease; osteoporosis; osteoarthritis; Alzheimer's disease; Parkinson's disease; Huntington's disease; infection; cardiac disease; stroke; sepsis; inflammatory bowel disease; contraceptive; immunogen; food supplement.

SO Homo sapiens.

XX
XX

PN WO200175068-

XX

PD 11-OCT-2001.

XX

22-MAR-2001; 2001WO-US0936

XX
XX

30-MAR-2000; 2000US-0539330.

04-DEC-2000; 2000US-0729674.

PA (GEMV) GENETICS TMO

FA (GEM) / GENETICS
XX

Jacobs K, McCoy JM, Lavallie E, Collins-racie LA, Evans C;
 Treacy M, Agostino MJ, Steininger RJ, Spaulding V, Wong GG;
 Clark H, Fectel K, Metzberg D;
 WPI: 2001-639363/73.
 P-PSDB: AAU38983.
 Secreted human proteins, useful as vaccine for treating various
 diseases such as autoimmune disorders (e.g. multiple sclerosis), and
 nervous system disorders (e.g. stroke) -

1000

Claim 1; Page 456-457; 61pp; English.

The invention relates to novel human secreted proteins, the nucleic acids encoding them. The protein may exhibit cytokine, cell proliferation or cell differentiation activity or may induce production of other cytokines in certain cell populations and may exhibit immune stimulating or immune suppressing activity, which is useful for the treatment of various immune deficiencies and disorders e.g. severe combined immunodeficiency (SCID), autoimmune disorders e.g. multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation. The proteins are also useful in the treatment of diseases and disorders including tissue, skin and organ transplantation and in graft-versus-host diseases (GVHD), in the induction of tumour immunity, myeloid or lymphoid cell deficiencies (GVHD), in the treatment of tumour repair, in the treatment of burns, incisions and ulcers; as well as in treatment of periodontal disease, osteoporosis or osteoarthritis, mediated by inflammatory processes, diseases of the peripheral nervous system, Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome, infections, infarction of cardiac and central nervous system vessel e.g. stroke, sepsis, inflammatory bowel disease, ulcers, bone regeneration. The protein, having active or inhibit-related activities is useful as a contraceptive based on the ability of inhibins to decrease fertility in


```
Db 3121 AAGTAGGATTTGGTGGACAGAGCGTCTGTGTGGTTGGATGGAGGTTCTGGCTGCA 3180
Qy 3181 AATCAAGTGGAGATTTCTCTAGGTCAGGTCTGTGACAGAGAGCTGACAGAGGATCTGAA 3240
Db 3181 AATCAAGTGGAGATTTCTCTAGGTCAGGTCTGTGACAGAGAGCTGACAGAGGATCTGAA 3240
Qy 3241 TGACATTTGGTTTATTTGTTGGGGTGTCTCTCAGAGGACCTGTGAAAGCCCTTTATCAGTC 3300
Db 3241 TGACATTTGGTTTATTTGTTGGGGTGTCTCTCAGAGGACCTGTGAAAGCCCTTTATCAGTC 3300
Qy 3301 ATTATTTGGCTGTGAGAAATTTCTCTGGAGTGTGGGTACATTTGAAGCAAGTGAATTC 3360
Db 3301 ATTATTTGGCTGTGAGAAATTTCTCTGGAGTGTGGGTACATTTGAAGCAAGTGAATTC 3360
Qy 3361 GTTGAGGGCAAGTCTCTGGAAGAGAGCTGTAGGATCTGGCAGCTACCAATCGGTGGTAG 3420
Db 3361 GTTGAGGGCAAGTCTCTGGAAGAGAGCTGTAGGATCTGGCAGCTACCAATCGGTGGTAG 3420
Qy 3421 TGTGTTGGGGTGGGGTCTCTGGGCACTGGCTGTGTGAAAGGATCTGGCAGGGCACCA 3480
Db 3421 TGTGTTGGGGTGGGGTCTCTGGGCACTGGCTGTGTGAAAGGATCTGGCAGGGCACCA 3480
Qy 3481 GCGCCCTCTACTGAACCATCAGCATCTCAGTGGCATTAAAGCCATGAGCTGAGAGGGC 3540
Db 3481 GCGCCCTCTACTGAACCATCAGCATCTCAGTGGCATTAAAGCCATGAGCTGAGAGGGC 3540
Qy 3541 CATGAGATTTCTCTGAGTATTACTGAGAGCAACAGAAAGAGCCATGATGAGAGGCC 3600
Db 3541 CATGAGATTTCTCTGAGTATTACTGAGAGCAACAGAAAGAGCCATGATGAGAGGCC 3600
Qy 3601 TTGGGCTCTCTGGGAATGGGAATCAGCCAAAGGACTGAGAAGGAGTACCTTAAGGTC 3660
Db 3601 TTGGGCTCTCTGGGAATGGGAATCAGCCAAAGGACTGAGAAGGAGTACCTTAAGGTC 3660
Qy 3661 AGAGAAACCAAGAGAGTGTGTGTTCTGAGAGCTGAGCTTTCTTTATCACTCATTC 3720
Db 3661 AGAGAAACCAAGAGAGTGTGTGTTCTGAGAGCTGAGCTTTCTTTATCACTCATTC 3720
Qy 3721 CCTTCTCCAAATTAAGCCACTGTGTAGTTGGGCCCCCTCAGAGGTTGAAGGCAAGAGGAGA 3780
Db 3721 CCTTCTCCAAATTAAGCCACTGTGTAGTTGGGCCCCCTCAGAGGTTGAAGGCAAGAGGAGA 3780
Qy 3781 AAGCAGAGCTTTGGGAACAGAGCTTTCTGCAATAGCTGCGAAGGAGGATTAAGGA 3840
Db 3781 AAGCAGAGCTTTGGGAACAGAGCTTTCTGCAATAGCTGCGAAGGAGGATTAAGGA 3840
Qy 3841 TAGAGTGTAAATAAAAAAAAAAAAAAAAAAAAA 3871
Db 3841 TAGAGTGTAAATAAAAAAAAAAAAAAAAAAAAA 3871
RESULT 3
ABA90876
ID ABA90876 standard; cDNA; 3871 BP.
XX ABA90876;
XX
XX
XX 14-FEB-2002 (first entry)
XX
XX Human polynucleotide SEQ ID NO 1.
XX
XX Human; clone bd306-7; clone yb8-1; ATCC number 98599; gene therapy;
XX immune disorder; bacterial infection; fungal infection; cancer; tumour;
XX autoimmune disorder; systemic lupus erythematosus; wound; ulcer; inhibin;
XX osteoporosis; osteoarthritis; nervous system disorder; neuropathy;
XX Alzheimer's disease; Parkinson's disease; Huntington's disease; activin;
XX haemophilia; cardiac infarction; stroke; sepsis; arthritis; vulnery;
XX ischaemia-reperfusion injury; inflammatory bowel disease; chemotactic;
XX Crohn's disease; cytostatic; anti-inflammatory; immunomodulator;
XX neuroprotective; haemostatic; thrombolytic; anti-inflammatory; ss.
XX Homo sapiens.
```

```
XX
PN
XX
PD
XX
PF
XX
XX 08-NOV-2001.
XX
XX 04-DEC-2000; 2000US-0729674.
XX
XX 10-APR-1997; 97US-126425P.
XX 04-DEC-1997; 97US-067454P.
XX 20-DEC-1997; 97US-068379P.
XX 02-JAN-1998; 98US-070346P.
XX 07-JAN-1998; 98US-070643P.
XX 08-JAN-1998; 98US-070755P.
XX 13-JAN-1998; 98US-071304P.
XX 22-JAN-1998; 98US-072134P.
XX 30-JAN-1998; 98US-073095P.
XX 18-FEB-1998; 98US-075038P.
XX 30-MAR-2000; 2000US-0539330.
XX 23-NOV-1998; 98US-0197886.
XX
XX (JACO/) JACOBS K.
XX (MCCO/) MCCOY J M.
XX (LAVA/) LAVALLIE E R.
XX (COLL/) COLLINS-RACIE L A.
XX (EVAN/) EVANS C.
XX (MERB/) MERBERG D.
XX (TREA/) TREACY M.
XX (AGOS/) AGOSTINO M J.
XX (STEL/) STEININGER R J.
XX (SPAU/) SPAULDING V.
XX (WONG/) WONG G G.
XX (CLAR/) CLARK H.
XX (FECH/) FECHTEL K.
XX
XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C,
XX Merberg D, Treacy M, Agostino MJ, Steininger RJ, Spaulding V;
XX Wong GG, Clark H, Fechtel K;
XX WPI; 2002-040725/05.
XX P-PSDB; ABB55698.
XX
XX New secreted proteins and encoding polynucleotides, useful in gene
XX therapies, particularly for preventing or treating autoimmune
XX disorders, cancer, graft-versus-host disease, wound, osteoporosis,
XX stroke or inflammations
XX
XX Claim 1; Page 175-177; 349pp; English.
XX
XX The invention relates to isolated polynucleotides (ABA90876-ABA90968 and
XX ABA90980) and encoded proteins (ABB55698-ABB58000), especially
XX polynucleotides SEQ ID NO 1 (ABA90876) and SEQ ID NO 19 (ABA90985) and
XX proteins SEQ ID NO 2 (ABB55698) and SEQ ID NO 20 (ABB55707) contained in
XX clones bd306-7 and yb8-1 respectively and the clones bd306-7 and yb8-1
XX are deposited with the American Type Culture Collection (ATCC) with
XX accession number 98599. The polynucleotides and encoded polypeptides have
XX cytostatic, anti-inflammatory, immunomodulator, vulnerary,
XX neuroprotective, activin, inhibin, chemotactic, haemostatic, thrombolytic
XX and anti-inflammatory activity and acting as cytokine modulators,
XX haematopoiesis regulators, tissue growth modulators and/or cadherin
XX suppressors. The polypeptides and polynucleotides are useful in gene
XX therapies, particularly for preventing, treating or ameliorating any of
XX the following diseases: immune deficiency and disorders; e.g. bacterial
XX or fungal infections, autoimmune disorders, cancer, systemic lupus
XX erythematosus or graft-versus-host disease; myeloid or lymphoid cell
XX deficiencies; wound, burns, incisions and ulcers, osteoporosis or
XX osteoarthritis; central and peripheral nervous system diseases and
XX neuropathies, e.g. Alzheimer's, Parkinson's disease, Huntington's
XX disease, amyotrophic lateral sclerosis or Shy-Drager syndrome;
XX haemophilia, cardiac infarction or stroke; inflammations, shock, sepsis
XX or systemic inflammatory response syndrome; ischaemia-reperfusion
XX injury, endotoxin lethality, arthritis, inflammatory bowel disease or
XX Crohn's disease; or tumours or cancers, pemphigus vulgaris or pemphigus
XX foliaceus.
```


XX 12-MAR-1999; 99US-0124270.
XX (HUMA-) HUMAN GENOME SCI INC.
XX (ROSE/) ROSEN C A.
XX Rosen CA, Ruben SM;
XX WPI; 2000-587513/55.
XX DR P-PSDB; AAB57022.
XX
XX Prostate cancer associated gene sequences, referred to as prostate
XX cancer antigens, useful for treatment, prevention, and diagnosis of
XX disorders such as prostate cancer -
XX
XX Claim 1; Page 1097-1098; 238pp; English.
XX
XX AAF1566 to AAF16505 encode the human prostate cancer associated
XX proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
XX The prostate cancer antigens can have neuroprotective, cytostatic,
XX cardioactive, immunomodulatory, muscular, vulnerary, gastrointestinal,
XX nephrotropic, antiinfective, gynecological and antibacterial activities,
XX and can be used in gene therapy. The prostate cancer antigen
XX polynucleotides may be used for detection of prostate cancer, chromosome
XX identification, as chromosome markers, and for numerous other diagnostic
XX or research purposes. The prostate cancer antigens may be used to treat
XX disorders such as neural, immune, muscular, reproductive,
XX gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
XX disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
XX AAB57303 represent sequences used in the exemplification of the present
XX invention.
XX
XX Sequence 2549 BP; 667 A; 519 C; 683 G; 673 T; 7 other;
XX
XX Query Match 53.0%; Score 2051.2; DB 21; Length 2549;
XX Best Local Similarity 99.2%; Pred. No. 0;
XX Matches 2077; Conservative 6; Mismatches 4; Indels 7; Gaps 2;
XX
XX 1763 CATGCCGGCCTCTTCTCACCTTTACACCTGTCTTCTTATCTTCTCACATCTGTTTCA 1822
XX 206 CTGCGCCGGCCTCTTCTCACCTTTACACCTGTCTTCTTATCTTCTCACATCTGTTTCA 265
XX 1823 CTTTCATCCCT 1882
XX 266 CTTTCATCCCT 325
XX 1893 TTACATTTTGGTTTGAAGGCGAGTCTTCTCTGCTGTTTCTTCTGTTTCTTCTGTTT 1942
XX 326 TTACCATTTTGGTTTGAAGGCGAGTCTTCTCTGCTGTTTCTTCTGTTTCTTCTGTTT 385
XX 1943 TCAGTATTATTTTAAATAAGAAAAACATTCCTAGAGATGATAATTTGTAAAAACCTCC 2002
XX 386 TCAGTATTATTTTAAATAAGAAAAACATTCCTAGAGATGATAATTTGTAAAAACCTCC 445
XX 2003 TTGGCTTTATTTGCTTTTTCAGATTGTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2062
XX 446 TTGGCTTTATTTGCTTTTTCAGATTGTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 505
XX 2063 GAAGACATAGGCTAAATTTCTCCAGCTTCAATAGTCTTCACTTGTGCTGACTTGTACC 2122
XX 506 GAAGACATAGGCTAAATTTCTCCAGCTTCAATAGTCTTCACTTGTGCTGACTTGTACC 565
XX 2123 AATTTAGCCACCTGTAAGAACAGTTGAGTAGAGTGTAGAGTGTAGAGTGTAGAGTGTAG 2182
XX 566 AATTTAGCCACCTGTAAGAACAGTTGAGTAGAGTGTAGAGTGTAGAGTGTAGAGTGTAG 625
XX 2183 TTGGCCCACTTTGTCATCTCAAAATTTACACGGTTGGCCGATCCCATTTTGGAGCAATG 2242
XX 626 TTGGCCCACTTTGTCATCTCAAAATTTACACGGTTGGCCGATCCCATTTTGGAGCAATG 685
XX 2243 CTTAGTTATAGTCTCCAGTTGGAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2302
XX 686 CTTAGTTATAGTCTCCAGTTGGAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 745

QY 2303 TTCTTTTCAAGTAAATATTTTATTTAGTACTACTGTGTGCTAGGCAATTCACCTGG-----G 2356
Db 746 TTCTTTTCAAGTAAATATTTTATTTAGTACTACTGTGTGCTAGGCAATTCACCTGGAACTAG 805
QY 2357 AACTAGAGATACCTTCACAGAAATACAGGAAAGTTCCCTGTGTCTCATGAGCTTACATTC 2416
Db 806 AACTAGAGATACCTTCACAGAAATACAGGAAAGTTCCCTGTGTCTCATGAGCTTACATTC 865
QY 2417 TACAGGGAAGAAAGATAGCAATACATAGGAATAAATATATATACAAAGGTATCATGTAGTG 2476
Db 866 TACAGGGAAGAAAGATAGCAATACATAGGAATAAATATATATACAAAGGTATCATGTAGTG 925
QY 2477 ATATTGCTGTGAGAAAAAATAAAGCAGGAGAGGAGTAAAGAAATCCTGAGATGAGGCT 2536
Db 926 ATATTGCTGTGAGAAAAAATAAAGCAGGAGAGGAGTAAAGAAATCCTGAGATGAGGCT 985
QY 2537 GCAGTTTTAAATGGGGCCTCACTGGGAATGTGACGTTGAGCAGAGACGTTAGGGAAGTGG 2596
Db 986 GCAGTTTTAAATGGGGCCTCACTGGGAATGTGACGTTGAGCAGAGACGTTAGGGAAGTGG 1045
QY 2597 ATCTGACAAAGGATTCAGGAGAGAGAAACAAGATGTGCACTGCCCCCAAGTGAAGT 2656
Db 1046 ATCTGACAAAGGATTCAGGAGAGAGAAACAAGATGTGCACTGCCCCCAAGTGAAGT 1105
QY 2657 TGCTCTACGTGTCAGGAAAGAGCAGGAGACCAAGCAGAGTCGTGGCAGGGGTAGAAT 2716
Db 1106 TGCTCTACGTGTCAGGAAAGAGCAGGAGACCAAGCAGAGTCGTGGCAGGGGTAGAAT 1165
QY 2717 GGAAGGAGAGGCGCTGGGAGAGACAGGTGTGTGGAGGCGCTTGGCTTCTGCTAAGTGAGA 2776
Db 1166 GGAAGGAGAGGCGCTGGGAGAGACAGGTGTGTGGAGGCGCTTGGCTTCTGCTAAGTGAGA 1225
QY 2777 TGGAAACCACTGGAGGGTGTGAAACAGAGAGTGCCTTTGATTTATATATTTTGAAGGG 2836
Db 1226 TGGAAACCACTGGAGGGTGTGAAACAGAGAGTGCCTTTGATTTATATATTTTGAAGGG 1285
QY 2837 TCATTCTAGTGCATATTTGTGAAACACTTTAGTGACAAAGGCGCAGAGAGAGAGAGAG 2896
Db 1286 TCATTCTAGTGCATATTTGTGAAACACTTTAGTGACAAAGGCGCAGAGAGAGAGAGAG 1345
QY 2897 ACCTGTTAGGAAGCTACTGCAAGGTTCCAGGCTTGGGCGCTGGGCGCACACAGCAGTG 2956
Db 1346 ACCTGTTAGGAAGCTACTGCAAGGTTCCAGGCTTGGGCGCTGGGCGCACACAGCAGTG 1405
QY 2957 GTCAATATCTAGATTTATTTTGAAGAGCCAAATAGGATTTGCTGAGATTTGAATGTG 3016
Db 1406 GTCAATATCTAGATTTATTTTGAAGAGCCAAATAGGATTTGCTGAGATTTGAATGTG 1465
QY 3017 GAGTGTAGGAAGAGAGTAAATGATGACATTAAGGTTTTTTGGCCTGAAATAGCAGGAA 3076
Db 1466 GAGTGTAGGAAGAGAGTAAATGATGACATTAAGGTTTTTTGGCCTGAAATAGCAGGAA 1525
QY 3077 AGATGGAGTTACCAAGTACTGAAATAGGAGAGGAGGCTGGGTAGTAAAGAAATTTGGT 3136
Db 1526 AGATGGAGTTACCAAGTACTGAAATAGGAGAGGAGGCTGGGTAGTAAAGAAATTTGGT 1585
QY 3137 GCAAGCAGGCTGTCTGTGTTGGAAATGGAGGTTCTGGCTGCAAAATCAAGATGGAGA-T 3195
Db 1586 GCAAGCAGGCTGTCTGTGTTGGAAATGGAGGTTCTGGCTGCAAAATCAAGATGGAGAGT 1645
QY 3196 TCTCTCAGGTCAGGTCGTGAGCAGAGCTGAGACAGGATCTGAATGCACTTTGGTTTATT 3255
Db 1646 TCTCTCAGGTCAGGTCGTGAGCAGAGCTGAGACAGGATCTGAATGCACTTTGGTTTATT 1705
QY 3256 GTTGGGGGTCTCTCAGAGAGAAACCTGTGTGAAAGCCTTTATCAGTCAATTTATTGGCTGTGA 3315
Db 1706 GTTGGGGGTCTCTCAGAGAGAAACCTGTGTGAAAGCCTTTATCAGTCAATTTATTGGCTGTGA 1765
QY 3316 GAGATTTCTCTGGAGTGTGGTACATTTTGAAGCAGTCACTTCACTTGGAGGCAAGTCT 3375
Db 1766 GAGATTTCTCTGGAGTGTGGTACATTTTGAAGCAGTCACTTCACTTGGAGGCAAGTCT 1825

684 AGCTGTTATCCCGAGAGTGGGGCATAAAGCGTGAGAACTGGAGAACATCCATCACA 743
Db
661 AGCTGTTATCCCGAGAGTGGGGCATAAAGCGTGAGAACTGGAGAACATCCATCACA 720
Qy
744 GACATGGAGTGGGCTTCCGAGCTTCTTCAAGGCTTTTACAGAGTGTCTCGAGGCTGC 803
Db
721 GACATGGAGTGGGCTTCCGAGCTTCTTCAAGGCTTTTACAGAGTGTCTCGAGGCTGC 780
Qy
804 GAGGTTCCAGGAGATCAAGGATCTTCAAGGATTTTACCTTTCCATAGCAGATCATAT 863
Db
781 GAGGTTCCAGGAGATCAAGGATCTTCAAGGATTTTACCTTTCCATAGCAGATCATAT 840
Qy
864 GTAGAAGTCTGGAATGCAAAATACAGTGTGAAGAGAACCTCACCCAGTTATAGGAGGC 923
Db
841 GTAGAAGTCTGGAATGCAAAATACAGTGTGAAGAGAACCTCACCCAGTTATAGGAGGC 900
Qy
924 TATCCGTTGAGAAATTTGGCTTACCATGATCATTTACTGTGAGTTTGGCTTATTAAG 983
Db
901 TATCCGTTGAGAAATTTGGCTTACCATGATCATTTACTGTGAGTTTGGCTTATTAAG 960
Qy
984 TTGAACGACCTGAAGATGCAAGCCCTGTGCACTGAGCTATCTGCTTTTGTATCAGAT 1043
Db
961 TTGAACGACCTGAAGATGCAAGCCCTGTGCACTGAGCTATCTGCTTTTGTATCAGAT 1020
Qy
1044 GACAGGTCTGACAGAGAACTGGTGTATTAACAGTACACAGGAGACATTTGGGCGCTC 1103
Db
1021 GACAGGTCTGACAGAGAACTGGTGTATTAACAGTACACAGGAGACATTTGGGCGCTC 1080
Qy
1104 TCGGATGAGCACTTCCAGCCAGACCTGAGCAGTTCAGTTCCTTTATGTGACCACTC 1163
Db
1081 TCAGATGAGCACTTCCAGCCAGACCTGAGCAGTTCAGTTCCTTTATGTGACCACTC 1140
Qy
1164 GACAGGAGCTGTATGACTTTGCTTAAGGAAATATATGATGATGATGAGGAGAGTT 1223
Db
1141 GACAGGAGCTGTATGACTTTGCTTAAGGAAATATATGATGATGATGAGGAGAGTT 1200
Qy
1224 GTGGAAATATGGATGACTTCTTGAACCTGGAGAGACCACTAGCCACAGCAACAA 1283
Db
1201 GTGGAAATATGGATGACTTCTTGAACCTGGAGAGACCACTAGCCACAGCAACAA 1260
Qy
1284 GAGACTTCTCTTGGCGTTCAGGAAACACAGATTTCTTTGCTTTTCCCAACAGCCAGG 1343
Db
1261 GAGACTTCTCTTGGCGTTCAGGAAACACAGATTTCTTTGCTTTTCCCAACAGCCAGG 1320
Qy
1344 CTGTTGATACCTCAGAGCTTCTCTTTACTCTCCAAAGTGAAGGGAGCCCGCTCTCT 1403
Db
1321 CTGTTGATACCTCAGAGCTTCTCTTTACTCTCCAAAGTGAAGGGAGCCCGCTCTCT 1380
Qy
1404 CTAACTGATGTCATCAGGGTGAAGCTGCTTCTTCTATCTTCACTGCGCACTCATG 1463
Db
1381 CTAACTGATGTCATCAGGGTGAAGCTGCTTCTTCTATCTTCACTGCGCACTCATG 1440
Qy
1464 TTCACACCTATCTTCTCAGCTTCTTTTGTGATGGAGTCTGCTCTCTTGGCCAGGCTG 1523
Db
1441 TTCACACCTATCTTCTCAGCTTCTTTTGTGATGGAGTCTGCTCTCTTGGCCAGGCTG 1500
Qy
1524 GAGTGCATGAGCACTTCTCAGCTCAGTGAACCTCCGCTCTTGGGTTCAAGCAATCT 1583
Db
1501 GAGTGCATGAGCACTTCTCAGCTCAGTGAACCTCCGCTCTTGGGTTCAAGCAATCT 1560
Qy
1584 GCTGCATCAGCTTCCGAGTACCTGGGATTAACAGGATGTCACACAGCCGCGCTAAT 1643
Db
1561 GCTGCATCAGCTTCCGAGTACCTGGGATTAACAGGATGTCACACAGCCGCGCTAAT 1620
Qy
1644 TTGTAATTTTAGTAGAGAGGGGTTTGGCAATGTTGGCAGGCTGTCTGAACTCTTGA 1703
Db
1621 TTGTAATTTTAGTAGAGAGGGGTTTGGCAATGTTGGCAGGCTGTCTGAACTCTTGA 1680
Qy
1704 CTTGAGATGATCATCTGCTTGGCTTCCACAGTGTGGATTAACAGGCTGTAGGCACC 1763
Db
1681 CTTGAGATGATCATCTGCTTGGCTTCCACAGTGTGGATTAACAGGCTGTAGGCACC 1740
Qy
1764 ATGCCGGGCTCTTCTCAGCTTTACACCTGTCTTCTTATCCTCACAATCTGTTTTCACAC 1823

1741 ATGCCGGGCTCTTCTCAGCTTTACACCTGTCTTATCTTATCTCCTCACAATCTGTTTTCACAC 1800
Qy
1824 CTTGATCCCTGCT 1883
Db
1801 CTTGATCCCTGCT 1860
Qy
1884 TACATATTTGGTTTGAAGGAGCT 1943
Db
1861 TACATATTTGGTTTGAAGGAGCT 1920
Qy
1944 CAGTATTTATTTTAAATAAGAAACATCT 1984
Db
1921 CAGTATTTATTTTAAATAAGAAACATCT 1961
RESULT 6
AAH31138
ID AAH31138 standard; cDNA; 1586 BP.
XX
AC AAH31138;
XX
DT 27-JUL-2001 (first entry)
XX
DE Human colon cancer cell line Km12L4-A cDNA library derived seq#1072.
XX
DE Human; diagnosis; colon cancer; cancer; malignant; chromosome mapping;
KW detection; colon cancer cell line Km12L4-A; ss.
XX
OS Homo sapiens.
XX
PN WO200018916-A2.
XX
PD 06-APR-2000.
XX
PF 23-SEP-1999; 99WO-US22226.
XX
PR 28-SEP-1998; 98US-0102161.
PR 28-SEP-1998; 98US-0102180.
PR 29-SEP-1998; 98US-0102380.
PR 08-OCT-1998; 98US-0103815.
PR 27-OCT-1998; 98US-0105877.
XX
(CHIR) CHIRON CORP.
(HYSE-) HYSEQ INC.
XX
PI Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
PI Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;
PI Lanson G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;
PI Leshkowitz D, Kita D, Garcia V, Jones LW, Stache-Grain B;
XX
WPI; 2000-293155/25.
XX
Polynucleotide library comprising 1079 defined sequences, useful in
the form of an array to detect cancer or susceptibility to cancer -
XX
Claim 1; Page 498-499; 502pp; English.
XX
The present invention describes a library of polynucleotides comprising
1079 nucleotide sequences (given in AAH30067 to AAH31145). Also described
are: (1) an isolated polynucleotide (I) having at least 90% identity to
one of the 1079 sequences; (2) a recombinant host cell containing (1);
(3) an isolated polypeptide (II) encoded by (1); (4) an antibody that
specifically binds to (II); (5) a vector comprising (1); and (6) a method
of detecting differentially expressed genes correlated with a cancerous
state of a mammalian cell comprising detecting a gene product encoded by
one of the 1079 sequences given in the specification. The polynucleotides
are used to monitor patients having (or susceptible) to cancer to detect
potentially malignant events at a molecular level before they are
detectable at a gross morphological level. The polynucleotides are also
useful for monitoring the efficacy of various therapies and preventive
interventions. Polynucleotide probes based on the disclosed sequences
are useful for chromosome mapping and detection of transcription levels.

[REDACTED]

DR WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.

PT Claim 1; Page 6191-6192; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV0010-ABV62213) of the specification or its complement. (I) is useful for:

CC (a) assessing whether a patient is afflicted with prostate cancer;

CC (b) monitoring the progression of prostate cancer in a patient;

CC (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient;

CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient;

CC (e) selecting a composition for inhibiting prostate cancer in a patient;

CC (f) assessing the prostate cell carcinogenic potential of a compound;

CC (g) determining whether prostate cancer has metastasized in a patient;

CC (h) assessing the aggressiveness or indolence of prostate cancer in a patient;

CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX Sequence 2499 BP; 589 A; 572 C; 701 G; 628 T; 9 other;

Query Match 27.4%; Score 1062; DB 23; Length 2499;

Best Local Similarity 99.4%; Pred. No. 4,7e-204;

Matches 1087; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

QY 2758 TGCTTCTCTCTAGTGTAGATGGAAACCACTGGA-CGGTTTGAACAGAGGAGTGCCTTGAT 2816

DB 29 TGTATCTCTCTAGTGTAGATGGAAACCACTGGAATGGTTTGAACAGAGGAGTGCCTTGAT 88

QY 2817 TGATTTATATTTTGAAGGGTCTATCTAGTCAATATTTGAAACCTTTAGTGGACAA 2876

DB 89 TGNATTTATTTTGCAGGGTCTATCTAGTCCATTTTGTGAAACCTTTAGTGGACAA 148

QY 2877 GGCCAGAGGAGAGGAGGAGAGCCTGTTAGGAGCTACTGCAAGTTTCAGGCTTGGGCT 2936

DB 149 GGCCAGAGGAGAGGAGGAGAGCCTGTTAGGAGCTACTGCAAGTTTCAGGCTTGGGCT 208

QY 2937 GGCCACAGCAACAGCAGTGGTCAATATCTAGATTTATTTTGAAGAGGCAATAGAT 2996

DB 209 GGCCACAGCAACAGCAGTGGTCAATATCTAGATTTATTTTGAAGAGGCAATAGAT 268

QY 2997 TTGCTGAGAGTTGAATGTTGAGTGTAAAGAGGAGGAGTAAATGATGACATTAAGTT 3056

DB 269 TTGCTGAGAGTTGAATGTTGAGTGTAAAGAGGAGGAGTAAATGATGACATTAAGTT 328

QY 3057 TTGGCTCAATAGCAGGAGGAGGAGTACAGTTACAGTTACTGAAATAGGAGGAGTGGCT 3116

DB 329 TTGGCTCAATAGCAGGAGGAGGAGTACAGTTACTGAAATAGGAGGAGTGGCT 388

QY 3117 GGTAAGTAAAGGATTTGGTCAAGAGGCTGTCTGTGTTGGAATGGGAGGTTCTGGC 3176

DB 389 GGTAAGTAAAGGATTTGGTCAAGAGGCTGTCTGTGTTGGAATGGGAGGTTCTGGC 448

QY 3177 TGCAATCAAGTGGAGAG-TTCTCTCAGTCAAGTCTGAGCAGAGCTCCAGAGCAGGAT 3235

DB 449 TGCAATCAAGTGGAGAGTTCTCTCAGTCAAGTCTGAGCAGAGCTCCAGAGCAGGAT 508

QY 3236 CTGAATGCACTTGGTTTATTTGTTGGGGTGTCTCAGAGGAACTGTGAAAGCTTTAT 3295

DB 509 CTGAATGCACTTGGTTTATTTGTTGGGGTGTCTCAGAGGAACTGTGAAAGCTTTAT 568

QY 3296 CAGTCATTTATTTGGTGTGAGAGTTCTCTGGAGTGGGTACATTTGAAGGCAAGTGA 3355

DB 569 CAGTCATTTATTTGGTGTGAGAGTTCTCTGGAGTGGGTACATTTGAAGGCAAGTGA 628

QY 3356 CTTCACTTGAAGGCAAGTCTCTGAAAGAGAGGCTGTAGGCATCTGGCAGCTACCAAGCT 3415

DB 629 CTTCACTTGAAGGCAAGTCTCTGAAAGAGAGGCTGTAGGCATCTGGCAGCTACCAAGCT 688

QY 3416 GGTAGTGTGTTGGGGTGGGGTCTCTGGGCACTGGCTGTGTGAAGGATCTGCGAGGCA 3475

DB 699 GGTAGTGTGTTGGGGTGGGGTCTCTGGGCACTGGCTGTGTGAAGGATCTGCGAGGCA 748

QY 3476 CCACAGCCCCCTACTGAACCATCAGCATGTCTAGTGGCATTTAAAGCCATGAGCTGGA 3535

DB 749 CCACAGCCCCCTACTGAACCATCAGCATGTCTAGTGGCATTTAAAGCCATGAGCTGGA 808

QY 3536 GGGGCCACTGAGATTGTCTCTGAGTATTACTTGAGAACCAACAGAAAGAGCCATGATGG 3595

DB 809 GGGGCCACTGAGATTGTCTCTGAGTATTACTTGAGAACCAACAGAAAGAGCCATGATGG 868

QY 3596 AGCCCTTGGGCTCTCTGGAAATGGAAATCAGCAAGGACTGAGAAAGAGTACCTTA 3655

DB 869 AGCCCTTGGGCTCTCTGGAAATGGAAATCAGCAAGGACTGAGAAAGAGTACCTTA 928

QY 3656 AGGTGAGAGAAACCAAGAGAGTGTGTCTGGAAGCTGAGCTTCTTTTCAACCT 3715

DB 929 AGGTGAGAGAAACCAAGAGAGTGTGTCTGGAAGCTGAGCTTCTTTTCAACCT 988

QY 3716 CATTCCTTCTCCAAATAAGCCACTTGTAGTTGGCCCTCCAGGTTGAAAGGCAAGA 3775

DB 989 CATTCCTTCTCCAAATAAGCCACTTGTAGTTGGCCCTCCAGGTTGAAAGGCAAGA 1048

QY 3776 GGAGAAAGGCACAGCGTTTGGAAACCAAGACTTTTCTGCAATAGCTGGAGGAAATAA 3835

DB 1049 GGAGAAAGGCACAGCGTTTGGAAACCAAGACTTTTCTGCAATAGCTGGAGGAAATAA 1108

QY 3836 AAGGATAGAGTGT 3849

DB 1109 AAGGATAGAGTGT 1122

RESULT 8

ABV30175

ID ABV30175 standard; cDNA; 2499 BP.

XX AC ABV30175;

XX DT 16-SEP-2002 (first entry)

XX DE Human prostate expression marker cDNA 30166.

XX DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX KW Pharmacogenomic marker; gene; ss.

XX OS Homo sapiens.

XX PN WO200160860-A2.

XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001WO-US05171.

XX PR 17-FEB-2000; 2000US-183319P.

XX PR 16-MAR-2000; 2000US-189862P.

XX PR 25-MAY-2000; 2000US-207454P.

XX PR 09-JUN-2000; 2000US-211314P.

XX PR 18-JUL-2000; 2000US-219007P.

XX PR 13-DEC-2000; 2000US-255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Endege WO, Monahan JE;

XX DR WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.

XX Claim 1; Page 6535; 11750pp; English.

CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

XX SQ Sequence 818 BP; 159 A; 271 C; 236 G; 148 T; 4 other;

Query Match 18.0%; Score 695.6; DB 22; Length 818;
 Best Local Similarity 96.7%; Pred. No. 1.9e-130;
 Matches 739; Conservative 2; Mismatches 18; Indels 5; Gaps 3;

QY 24 CTTCTTCGCTCCCTTCCTTCCTTCCTTCGCGGGCGGATGAGCGGGCGCGGGG 83
 DB 1 CTTCTTCGCTCCCTTCCTTCCTTCCTTCGCGGGCGGATGAGCGGGCGCGGGG 60
 QY 84 GCGGGCGGCTGCTAGCGCTCTGTGCTGCTGCTGCGGCGCGGGCGGCCAA 143
 DB 61 GCGGGCGGCTGCTAGCGCTCTGTGCTGCTGCTGCGGCGCGGGCGGCCAA 120
 QY 144 TAGCAACGCTACAGCTTCGCGAGCTTCCTACGGAAGAGCTGATGCGCTCGAGTCGGCC 203
 DB 121 TAGCAACGCTACAGCTTCGCGAGCTTCCTACGGAAGAGCTGATGCGCTCGAGTCGGCC 180
 QY 204 TACCGGACGCGCTGCAAGTACAGCGGGCGGAGCACTGGCGCGAGAGCTGCGCTACCTG 263
 DB 181 TACCGGACGCGCTGCGCAAGTACAGCGGGCGGAGCACTGGCGCGAGAGCTGCGCTACCTG 240
 QY 264 GAGATCAGCTGCGGCTGCAAGCTGCTGCTGCGGAGAGAGGCTTCTGCCACCGCAAC 323
 DB 241 GAGATCAGCTGCGGCTGCAAGCTGCTGCTGCGGAGAGAGGCTTCTGCCACCGCAAC 300
 QY 324 TGCAGCGCGGCGCGAGCGCGCGCGCGCGCGCTGCGCGAGCTATCCGAGCTGCGC 383
 DB 301 TGCAGCGCGGCGCGAGCGCGCGCGCGCGCGCTGCGCGAGCTATCCGAGCTGCGC 360
 QY 384 CTCCTTCGCGGCTGCTGCGCGCGCGCGCGCGCTGCTCAAGCGCTGCAAGCGGCTGCCA 443
 DB 361 CTCCTTCGCGGCTGCTGCGCGCGCGCGCGCGCTGCTCAAGCGCTGCAAGCGGCTGCCA 420
 QY 444 GCTTTCGCCAGTCCCGAGCGCGCGAGGTGCTGCGGAGCTTCCAGCGCGCGAGCGCC 503
 DB 421 GCTTTCGCCAGTCCCGAGCGCGCGAGGTGCTGCGGAGCTTCCAGCGCGCGAGCGCC 480
 QY 504 TACAAGTTCCTGAGTTCGCTTACTTCAAGCAAAATATCTGCCAAAGCATTCGCGCT 563
 DB 481 TACAAGTTCCTGAGTTCGCTTACTTCAAGCAAAATATCTGCCAAAGCATTCGCGCT 540
 QY 564 GCTCACACTTCTTACTCAAGCATCTGATGACGAAATGATGAGAGGACATGGCATAT 623
 DB 541 GCTCACACTTCTTACTCAAGCATCTGATGACGAAATGATGAGAGGACATGGCATAT 600
 QY 624 TATAAGAGCTGCTGCTGCTGCGAGGATACATTAAGAGCTGGAACCAAGT--CATATG 681
 DB 601 TATAAGAGCTGCTGCTGCTGCGAGGATACATTAAGAGCTGGAACCAAGT--CATATG 659
 QY 682 AAAGGCTGTTTATCGGAGCTGCGGCGATACACGCTGAGACTGGAGAACATCCATCA 741
 DB 660 AAAGGCTGTTTATCGGAGCTGCGGCGATACACGCTGAGACTGGAGAACATCCATCA 719
 QY 742 CAGACATGAGCTGCGCC--TTCCCGACTTCTTCAAGAGCTTTT 783

DB 720 CAGACATGGAAGTGGCCCTTCGGAATTTCTTCAAAAGNCTTT 763
 RESULT 10
 ABQ56517
 ID ABQ56517 standard; cDNA; 660 BP.
 XX ABQ56517;
 AC ABQ56517;
 XX 02-AUG-2002 (first entry)
 XX Human colon cancer related nucleotide sequence SEQ ID NO:212.
 XX Human; colon cancer; cancer; tissue profiling; forensic; mapping;
 KW genetic analysis; diagnostic; antisense therapy; gene; ss.
 XX Homo sapiens.
 XX WO200229086-A2.
 XX 11-APR-2002.
 XX 02-OCT-2001; 2001WO-US30732.
 XX 02-OCT-2000; 2000US-237271P.
 XX (FARB) BAYER CORP.
 PA Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;
 PI Thiagalingam A, Lewis ME;
 XX WPI; 2002-426115/45.
 XX New isolated nucleic acid that is differentially expressed in cancer
 PT tissues useful for determining the presence of colon cancer in a cell
 PT or tissue type, and in antisense therapy
 XX Claim 1; Fig 1; 796pp; English.
 CC ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
 CC expressed in cancer tissues. ABQ78993 to ABQ79004 represent proteins
 CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
 CC used in antisense therapy. An antibody immunoreactive with a polypeptide
 CC encoded by (I) is useful for detecting cancer in a patient sample, and
 CC for detecting the presence or absence of a polynucleotide encoded by a
 CC nucleic acid which hybridises to (I) in a cell. A probe/primer derived
 CC from (I) can be used for determining the presence of a nucleic acid which
 CC hybridises to (I), and for determining the phenotype of cells in a sample
 CC of cells from a patient. (I) is useful for determining the presence of
 CC colon cancer in a cell or tissue type, for determining the presence or
 CC state of other type of cancer, in antisense therapy, to generate
 CC macroarrays on a solid surface, to identify a chromosome on which the
 CC corresponding gene resides, and in tissue profiling, forensics, genetic
 CC analysis, mapping and diagnostic applications. (I) can be used to raise
 CC antibodies, and to screen for peptide analogues and antagonists.
 XX Sequence 660 BP; 180 A; 100 C; 211 G; 166 T; 3 other;

Query Match 15.2%; Score 587.4; DB 24; Length 660;
 Best Local Similarity 97.4%; Pred. No. 1.1e-108;
 Matches 627; Conservative 0; Mismatches 14; Indels 3; Gaps 3;

QY 2699 CGTGGGCGAGGGGTAGTAATGGAAGAGAGCGCGCTGGGAGGACAGGTGTGAGGGCTT 2758
 DB 2 CGCGGGCGAGGGGTAGTAATGGAAGAGAGCGCGCTGGGAGGAGACAGGTGTGAGGGCTT 61
 QY 2759 GGCTTCTGCTAAGTGTAGATGGGAACCACTGGAGGGTTTGAACAGAGGAGTGCCTTATTG 2818
 DB 62 GGCTTCTGCTAAGTGTAGATGGGAACCACTGGAGGGTTTGAACAGAGGAGTGCCTTATTG 121
 QY 2819 ATTATATTTTGAAGGTCATTCTAGCTGCAATATTGTGAAAAAATTGTAGTGAACAAG 2878
 DB 122 ATTATATTTTGAAGGTCATTCTAGCTGCAATATTGTGAAAAAATTGTAGTGAACAAG 181

PD 13-DEC-2001.
 XX 30-MAY-2001; 2001WO-US10838.
 XX 05-JUN-2000; 2000US-209473P.
 PR 05-JUN-2000; 2000US-209531P.
 PR 18-SEP-2000; 2000US-233133P.
 PR 18-SEP-2000; 2000US-233617P.
 PR 20-SEP-2000; 2000US-234009P.
 PR 20-SEP-2000; 2000US-234034P.
 PR 20-SEP-2000; 2000US-234052P.
 PR 22-SEP-2000; 2000US-234509P.
 PR 22-SEP-2000; 2000US-234567P.
 PR 22-SEP-2000; 2000US-234923P.
 PR 25-SEP-2000; 2000US-234924P.
 PR 25-SEP-2000; 2000US-235077P.
 PR 25-SEP-2000; 2000US-235082P.
 PR 25-SEP-2000; 2000US-235134P.
 PR 25-SEP-2000; 2000US-235280P.
 PR 26-SEP-2000; 2000US-235637P.
 PR 26-SEP-2000; 2000US-235639P.
 PR 27-SEP-2000; 2000US-235711P.
 PR 27-SEP-2000; 2000US-235720P.
 PR 27-SEP-2000; 2000US-235840P.
 PR 28-SEP-2000; 2000US-236028P.
 PR 28-SEP-2000; 2000US-236032P.
 PR 28-SEP-2000; 2000US-236033P.
 PR 28-SEP-2000; 2000US-236034P.
 PR 28-SEP-2000; 2000US-236109P.
 PR 28-SEP-2000; 2000US-236111P.
 PR 29-SEP-2000; 2000US-236842P.
 PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 PA (AVAL-) AVALON PHARM.
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 XX WPI; 2002-188264/24.
 PT Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set
 XX Claim 1; SEQ ID 5999; 44pp; English.
 PS The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABLE1664
 CC to AB170110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC

CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilm's tumour.
 XX
 SQ Sequence 2347 BP; 517 A; 689 C; 733 G; 408 T; 0 other;
 Query Match 12.0%; Score 465.8; DB 24; Length 2347;
 Best Local Similarity 64.2%; Pred. No. 4.2e-84;
 Matches 752; Conservative 2; Mismatches 399; Indels 18; Gaps 3;
 QY 76 GCCGGGGCGCGCGCTGTAGCGCTGCTGTGCTGCGCTGCGCGCTGCGCGCGCGC 135
 DB 1 GCGCGCGCGCATGGCTCGGGTGGCGGCTGCTGTGCTGTGCTGTGCGGCGCGCG 60
 QY 136 GCGCCCAATACGAACGCTACAGCTTCGCGAGTTCCACGGGACGAGCTGTATGCGCTCG 195
 DB 61 GGGCGCAGTACGAGAAGTACAGCTTCGCGGGCTTCGCGCCGAGGACCTGTATGCGCTGG 120
 QY 196 AGTCGGCTTACCGGCACGCGCTGGACAGTACAGCGGCGGACCTGGGCGGAGAGCTKG 255
 DB 121 CCGCGCGCTACGGGCACGCTCTGGAGACATACAGGGAGAGCTGGGCGGAGCGGCGC 180
 QY 256 GCTACTGTGAGATACGCTTGCCTGCGCTTGTGCTGCGCACAGCGAGGCTTCTGCC 315
 DB 181 GCTACTGTGAGCGCGCTGCGCTGCACTGCGCTCTTGCCTGCGCACAGCGAGGCTTCTGCC 240
 QY 316 ACGGCAACTGCAGCG-----CGGCGCGCAGCGCGGAGCGCGCGCGCTGCGCA 366
 DB 241 AGCCCAACTGCAGCGAG 300
 QY 367 GCTATCC-----CGAGTGGCGCTTCTGCGGGGCGCTGCTGCGCGCGCGCTGCGTCA 420
 DB 301 ACGAGTGGCGCTGCGAGCTGCGGCTCTTGGCGCGCTGCTGCGAGCGCGCTGCGTGC 360
 QY 421 ACGGCTGCGAGCGCGCTGCGAGCTTCCGCGCATCCGCGCGCGCGCGCGCGCGAGTGTGG 480
 DB 361 GCGCTGCAAGCGGCGCTGCGCGCTTCCAGGTGCGCTTCCGCGCGCGCGCGCGCGTGC 420
 QY 481 CGGACTTCCAGCGCGCGCGCGCGCTTCCAGGTGCGCTTCCAGGTGCGCTTCCAGGTGCGCTT 540
 DB 421 GTGACTTCCAGCGCGCGCTGCGCTTCCAGGTGCGCTTCCAGGTGCGCTTCCAGGTGCGCTT 480
 QY 541 ATCTCCCAAGCGCATCCGCGCTGCTACACCTTTCTTACTGAAGCATCTCTGATGCGAAA 600
 DB 481 GCGCTGGAAGCGGCTGCGCGCGCGCTTCCAGGTGCGCTTCCAGGTGCGCTTCCAGGTGCGCTT 540
 QY 601 TGATGAAGAGGAACATGCGCATATTATTA--GACCTGCTGCTGCTGCGGAGGACTACATTA 657
 DB 541 TGACCGCGCAAGTATCTCAACTACTATCAGGGGATGCTGCGCGCGCGCGCGCGCGCTCA 600
 QY 658 AAGACTCGAAACCAAGTATATGAAGCGCTTTCATCCGAGCGAGTGGCGCGCATACACG 717
 DB 601 CGGACTAGAGCGCGCGCGCTTACGAGCGCGCTTCCCGCGCGCTGCTGAGCTTCAACACT 660
 QY 718 GTGAGACTGAGAACATCCATCAGACATGAGCTGCGCGCTTCCCGACTTCTTCAAG 777
 DB 661 GCGGGGATTCGCGAGCAGCAGCGGAGCATGAGGAGCGCGCTTCTGCTGAGTACCTGCGAG 720
 QY 778 CTTTTCAGAGTGTCTGCGAGCGCTTCCAGGGGTTCCAGGGAGATCAAGACTTCAAGGATT 837
 DB 721 TCTTTCGCGCGTCTGCGCGCTGCTGAGGGGCGCGCTGAGCGGCTTCAAGACT 780
 QY 838 TCTACCTTTCATAGCAGATCATTTAGAGTTTGGAAATGGAATGCAAAATACAGTGAAG 897
 DB 781 TCTACCGCGCGCATAGCAGATCTCTTTGACAGTCCCTGCGAGTCAAGGTGAGCTGTGAGG 840
 QY 898 AGAACCTCACCCAGTTATAGGAGGCTATCCGGTTGAGAAATTTGGCTTACCATGTATC 957
 DB 841 CCAATTTGACCCCAATGTTGGTGGCTTCTGTTGACAAATGTTGTTGGCCCATGTATCC 900
 QY 958 ATTAATCTGCGTTTGGCTTATTAAGTTGAACCGCTGAAGATGCGAGCGCGCTGTGCGAG 1017

Db 901 ACTACCTGCAGTTTGGCTACTATAGTTGATGATGCGCCAGGCTGCCGAGCGCG 960
QY 1018 TCAGCTATCTGCTCTTTGATCAGATGACAGTCTATGAGCAGAACCTGGTATTACC 1077
Db 961 CCAGCTACATGCTCTTCGACCCCAAGGACAGCGCTCATGCGAGCAGAACCTGGTATTACC 1020
QY 1078 AGTACCACAGGACACTTGGGGCTCTCGGATGAGCAGCTCCAGCCACACCTGAGCAG 1137
Db 1021 GGTTCACCGGGCTGCTGCGGCTGAGAGAGGAGACTTCAGCCCCGGAGGAGGCA 1080
QY 1138 TTGAGTCTTTTAATGAGCACACTCCAGAGGAGCTGTATGACTTTGTAAGGAAATA 1197
Db 1081 TGCTCTACCAACACAGACCGCGAGCTGCGGAGCTGCTGGAGTTTCAACCCACATGTACC 1140
QY 1198 TAATGATGATGATGAGGAGAGTGTGGA 1228
Db 1141 TGCAGTCAGATGATGAGAGCTGAGGA 1171

RESULT 13

AAH30585
ID AAH30585 standard; cDNA; 408 BP.

AC AAH30585;

XX 27-JUL-2001 (first entry)

DE Human colon cancer cell line Kml2L4-A cDNA library derived sequence #519.

XX Human; diagnosis; colon cancer; cancer; malignant; chromosome mapping;
KW detection; colon cancer cell line Kml2L4-A; ss.

XX Homo sapiens.

XX WO200018916-A2.

XX 06-APR-2000.

XX 23-SEP-1999; 99WO-US222226.

XX 28-SEP-1998; 98US-0102161.

XX 28-SEP-1998; 98US-0102180.

XX 29-SEP-1998; 98US-0102380.

XX 08-OCT-1998; 98US-0103815.

XX 27-OCT-1998; 98US-0105877.

XX (CHIR) CHIRON CORP.

PA (HYSE-) HYSEQ INC.

XX Williams LT, Escobedo J, Innis WA, Garcia PD, Sudduth-Klinger J;
PI Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;
PI Lamon G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;
PI Leshkowitz D, Kita D, Garcia V, Jones LW, Stache-Crain B;
XX WPI; 2000-293155/25.

XX Polynucleotide library comprising 1079 defined sequences, useful in
PT the form of an array to detect cancer or susceptibility to cancer -
XX Claim 1; Page 333; 502pp; English.

XX The present invention describes a library of polynucleotides comprising
CC 1079 nucleotide sequences (given in AAH30067 to AAH31145). Also described
CC are: (1) an isolated polynucleotide (I) having at least 90% identity to
CC one of the 1079 sequences; (2) a recombinant host cell containing (I);
CC (3) an isolated polypeptide (II) encoded by (I); (4) an antibody that
CC specifically binds to (II); (5) a vector comprising (I); and (6) a method
CC of detecting differentially expressed genes correlated with a cancerous
CC state of a mammalian cell comprising detecting a gene product encoded by
CC 65 of the 1079 sequences given in the specification. The polynucleotides
CC are used to monitor patients having (or susceptible) to cancer to detect
CC potentially malignant events at a molecular level before they are
CC detectable at a gross morphological level. The polynucleotides are also

CC useful for monitoring the efficacy of various therapies and preventive
CC interventions. Polynucleotide probes based on the disclosed sequences
CC are useful for chromosome mapping and detection of transcription levels.
CC The 1079 polynucleotide sequences were derived from a human colon cancer
CC cell line Kml2L4-A cDNA library.

XX Sequence 408 BP; 117 A; 96 C; 98 G; 97 T; 0 other;

Query Match 10.5%; Score 406.4; DB 21; Length 408;

Best Local Similarity 99.8%; Pred. No. 2.5e-72;

Matches 407; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 559 CCGTGTCTCACACCTTTCTACTGAAGCATCCTGATGACGAAATGATGAAGGAACATGG 618

Db 1 CCGTGTCTCACACCTTTCTACTGAAGCATCCTGATGACGAAATGATGAAGGAACATGG 60

QY 619 CATATTATAAGAGCGCTGCGGAGGACTTAAAGACCTGGAACCAAGTCAT 678

Db 61 CATATTATAAGAGCGCTGCGGAGGACTTAAAGACCTGGAACCAAGTCAT 120

QY 679 ATGAAAGCCTGTTTCATCCGAGCAGTGGGGCATACACGGTGAGAACTGGAGAACATCCA 738

Db 121 ATGAAAGCCTGTTTCATCCGAGCAGTGGGGCATACACGGTGAGAACTGGAGAACATCCA 180

QY 739 TCACAGACATGGAGCTGGCCCTTCCCGACTTCTTCAAAGCCTTTTACGAGTGTCTGCAG 798

Db 181 TCACAGACATGGAGCTGGCCCTTCCCGACTTCTTCAAAGCCTTTTACGAGTGTCTGCAG 240

QY 799 CTTGCGAGGGTTCCAGGGAGATCAAGGACTTCAAGGATTTCTACCTTTCCATAGCAGATC 858

Db 241 CTTGCGAGGGTTCCAGGGAGATCAAGGACTTCAAGGATTTCTACCTTTCCATAGCAGATC 300

QY 859 ATTATGTAGAAGTTCTGGAATGCAAAATACAGTGTGAAGAGACCTCACCCAGTTATAG 918

Db 301 ATTATGTAGAAGTTCTGGAATGCAAAATACAGTGTGAAGAGACCTCACCCAGTTATAG 360

QY 919 GAGGTATCCGTTGAGAAATTTGTGGTACCATTGATCATTTACTTTC 966

Db 361 GAGGTATCCGTTGAGAAATTTGTGGTACCATTGATCATTTACTTTC 408

RESULT 14

ABN94955/c

ID ABN94955 standard; DNA; 394 BP.

XX ABN94955;

XX 13-AUG-2002 (first entry)

XX Gene #1453 used to diagnose liver cancer.

XX Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW metastatic liver tumour; cytostatic; expression profile; disease state;
KW disease progression; drug toxicity; drug efficacy; drug metabolism.

XX Homo sapiens.

XX WO200229103-A2.

XX 11-APR-2002.

XX 02-OCT-2001; 2001WO-US30589.

XX 02-OCT-2000; 2000US-237054P.

XX (GENE-) GENE LOGIC INC.

XX Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;

XX WPI; 2002-426119/45.

XX Diagnosing and detecting the progression of liver cancer,
PT hepatocellular carcinoma or metastatic liver tumor in a patient,

PT involves detecting the level of expression of two or more genes in a
PT liver tissue sample -
XX
PS Claim 1; SEQ ID NO 1453; 298pp; English.
XX
CC The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumour in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytosolic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pot_sequences.
XX
SQ Sequence 394 BP; 143 A; 65 C; 104 G; 82 T; 0 other;

Query Match 10.1%; Score 392.4; DB 24; Length 394;
Best Local Similarity 99.7%; Pred. No. 1.6e-69; Mismatches 0; Gaps 0;
Matches 393; Conservative 0; Indels 0; Gaps 0;
QY 1709 GATGATCCATCTGCTTGGCTCCACAGTGTGGGATACAGGCGTGAGCCACCATGCC 1768
DB 394 GATGATCCATCTGCTTGGCTCCACAGTGTGGGATACAGGCGTGAGCCACCATGCC 335
QY 1769 CGGCTCTTCTACACCTTACACCTGCTCTTATCTCTCTCTCTCTCTCTCTCTCA 1828
DB 334 CGGCTCTTCTACACCTTACACCTGCTCTTATCTCTCTCTCTCTCTCTCTCTCA 275
QY 1829 TCCTGCTCA 1888
DB 274 TCCTGCTCA 215
QY 1889 TTTTGGTTGAAGGGAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1948
DB 214 TTTTGGTTGAAGGGAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 155
QY 1949 TTTTATTTTAAATAAGAAACATCTCTAGAGATGATATTTGTGAAACCTCTCTTGGC 2008
DB 154 TTTTATTTTAAATAAGAAACATCTCTAGAGATGATATTTGTGAAACCTCTCTTGGC 95
QY 2009 TTTTGGCTTTTCCAGATTTTAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2068
DB 94 TTTTGGCTTTTCCAGATTTTAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 35
QY 2069 ATAGGCTAATTTCTCCAGCTCAGATGCTCTT 2102
DB 34 ATAGGCTAATTTCTCCAGCTCAGATGCTCTT 1

RESULT 15
AAH50786
ID AAH50786 standard; cDNA; 401 BP.
XX
AC AAH50786;
XX
DT 23-AUG-2001 (first entry)
XX
XX Human tumour associated cDNA #115.
DE
XX Human; cancer specific gene expression; gene therapy;
KW age related differential expression; ss.
XX
OS Homo sapiens.
XX
PN WO200136685-A2.
XX

PD 25-MAY-2001.
XX
PF 17-NOV-2000; 2000WO-US31809.
XX
XX 17-NOV-1999; 99US-0166056.
PR 17-NOV-1999; 99US-0166106.
XX
PA (NYXI-) NYXIS NEURO THERAPIES INC.
XX
XX Kroes RA, Moskal JR, Yamamoto H;
PI
XX WPI; 2001-355647/37.
DR
XX Novel nucleic acid molecules differentially expressed in brain cancers,
PT useful for ascertaining propensity of cell for malignant phenotype or
PT ascertaining suitability of anti-neoplastic drug candidate -
XX
PS Claim 28; Page 53; 82pp; English.
XX
CC The present invention provides the sequences of 184 cDNA fragments which
CC are differentially expressed in cancer cell depending on the age of the
CC patient. They can be used to diagnose and identify treatments for
CC cancers, particularly brain cancers such as haemangioblastoma, teratoma,
CC haemangioma, glioblastoma, schwannoma, osteoma and pinealoma. The
CC present sequence is a cancer-associated cDNA of the invention.
XX
SQ Sequence 401 BP; 133 A; 72 C; 90 G; 106 T; 0 other;
Query Match 9.6%; Score 371.6; DB 22; Length 401;
Best Local Similarity 97.7%; Pred. No. 2.5e-65; Mismatches 9; Indels 0; Gaps 0;
Matches 377; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 2129 AGCACCACCTGAAACCAAGTTGAGTAGAGAGTGTAGAGTGCAGAGTGTGGCTTTTGGC 2188
DB 14 AGCACCACCTGAAACCAAGTTGAGTAGAGAGTGTAGAGTGCAGAGTGTGGCTTTTGGC 73
QY 2189 CCACCTTTGCTATCTCCAAAATTACAAACGTTTGGCCGATCCCATTTAGGACAATGCTTAGT 2248
DB 74 CCACCTTTGCTATCTCCAAAATTACAAACGTTTGGCCGATCCCATTTAGGACAATGCTTAGT 133
QY 2249 TATAAGTCTCCGAGTTGGAAAAGGAAAGCCAGAGCTGTCTAGTTTTCATTCTTTT 2308
DB 134 TATAAGTCTCCGAGTTGGAAAAGGAAAGCCAGAGCTGTCTAGTTTTCATTCTTTT 193
QY 2309 CAGTAAATATTTATTGTAGTACTACTGTGTGCTAGGCAATGACCTGGGAACTAGAGATAC 2368
DB 194 CAGTAAATATTTATTGTAGTACTACTGTGTGCTAGGCAATGACCTGGGAACTAGAGATAC 253
QY 2369 TTACAGAAATAACAGGGAAGTTCCTGTGCTCTATGGAGCTTACATTCTACAGGAGAAA 2428
DB 254 TTACAGAAATAACAGGGAAGTTCCTGTGCTCTATGGAGCTTACATTCTACAGGAGAAA 313
QY 2429 GAGATAGCCAATACATAGGAATAAATATATACAAAGTATCATGTAGTGATAATTGCTGTG 2488
DB 314 GAGATAGCCAATACATAGGAATAAATATATACAAAGTATCATGTAGTGATAATTGCTGTG 373
QY 2489 GAGAAAAATAAGCAGGGAGGGAGT 2514
DB 374 GAAAAATAAGGCTTTAAGGGCGAAT 399

Search completed: January 30, 2004, 08:32:11
Job time : 1013.41 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2004, 07:55:54 ; Search time 8124.86 Seconds
(without alignments)
11579.600 Million cell updates/sec

Title: US-09-729-674-1
Perfect score: 3871
Sequence: 1 ttctctctccctcccttt.....ataaaaaaaaaaaaaaaa 3871

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST.*

- 1: em_estba.*
- 2: em_esthm.*
- 3: em_estin.*
- 4: em_estnu.*
- 5: em_estov.*
- 6: em_estpl.*
- 7: em_estro.*
- 8: em_hic.*
- 9: gb_est1.*
- 10: gb_est2.*
- 11: gb_hic.*
- 12: gb_est3.*
- 13: gb_est4.*
- 14: gb_est5.*
- 15: em_estfun.*
- 16: em_estom.*
- 17: em_gss_hum.*
- 18: em_gss_inv.*
- 19: em_gss_pin.*
- 20: em_gss_vrt.*
- 21: em_gss_fun.*
- 22: em_gss_man.*
- 23: em_gss_mus.*
- 24: em_gss_pro.*
- 25: em_gss_red.*
- 26: em_gss_png.*
- 27: em_gss_vrl.*
- 28: gb_gss1.*
- 29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1072.2	27.7	1201	9	AL545483
C 2	1012.4	26.2	1200	9	AL571267
C 3	1004.8	26.0	1201	9	AL569764
C 4	985.2	25.5	1196	9	AL517454

5	979.6	25.3	1201	9	AL514488
6	969.8	25.1	1201	9	AL556976
7	964.6	24.9	1201	13	EX425905
C 8	954.8	24.7	1201	13	EX425904
9	954	24.6	1148	13	BU902192
C 10	947.2	24.5	1046	13	EX380662
11	945.8	24.4	1075	9	AL546910
12	934.6	24.2	1201	9	AK047506
13	933.6	24.2	1201	9	AL541167
C 14	924.6	23.9	1074	9	AL572160
15	924.4	23.9	1652	11	AK017797
16	923	23.8	1201	9	AL517455
C 17	922.8	23.8	1148	9	AL574681
C 18	918.4	23.7	1201	9	AL579089
19	913.4	23.6	1064	12	BM557305
20	896.6	23.2	1053	12	BM471592
21	891	23.0	1201	13	EX437872
C 22	881	22.8	1201	9	AL568770
23	878.4	22.7	1131	12	BM554453
24	873.4	22.6	1201	9	AL554750
25	866.8	22.4	1201	9	AL543090
26	862.4	22.3	1201	9	AL558186
C 27	859.4	22.2	1021	9	AL580676
28	858.8	22.2	1032	12	BM558598
29	857	22.1	939	13	EX386576
C 30	829	21.4	1201	9	AL514487
31	827.2	21.4	1152	9	AL550966
32	824.8	21.3	1040	9	AL558966
33	810.8	20.9	1021	12	BM807166
34	804.6	20.8	1201	13	EX387779
35	801	20.7	846	13	BU838140
36	788	20.4	961	12	BI255097
37	787.6	20.3	944	13	BU856662
38	784.4	20.3	1109	12	BM550012
39	777	20.1	827	12	BI760650
40	773.6	20.0	1083	12	BM920687
41	771.4	19.9	774	9	AL598447
42	770.6	19.9	854	12	BI860855
43	766.8	19.8	899	13	BQ437797
44	765.2	19.8	845	12	BI253401
45	759.6	19.6	933	12	BI858585

ALIGNMENTS

RESULT 1
AL545483
LOCUS
DEFINITION AL545483 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1015YK22 5-PRIME, mRNA sequence.
ACCESSION AL545483
VERSION AL545483.2 GI:31267318
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li W.B., Gruber C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 15, 2001 this sequence version replaced gi:12877964.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 2233.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0D1015BF11QPI&cluster=2233.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :


```
QY 738 ATCACAGACATGAGCTGGCCCTTCCCGACTTCTTCAAGCCCTTTTACGAGTGTCTCGCA 797
Db 1193 MTCCTTCAACAAATDGGKSTGCCCTCCCAATGCTCAAA-CTTTTACGAKTTTCTCGCA 1135
QY 798 GCTCGGAGGTTCCAGGAGATCAAGGACTTCAAGGATTTCTACCTTTCCATAGCAGAT 857
Db 1134 SCCK--CGRGGTCCAGGRARCAAGGACTTCAAGGATTTCT-MCTTTCCATAGCAGAY 1078
QY 858 CATTATGTAGAGTCTGGGAATCAAAATACAGTGTGAAGAGAACCTCACCCAGTTATA 917
Db 1077 AATTATGTAGAGTCTGGGAATCAAAATACAGTGTGAAGAGAACCTCACCCAGTTATA 1021
QY 918 GGAGGCTATCCGCTTGCAGAAATTTGGCTACCATGATCATTTACCTTTGAGCTTCCCTAT 977
Db 1020 GRA-GCTATCCGCTTGCAGAAATTTGGCTACCATGATCATTTACCTTTGAGCTTCCCTAT 963
QY 978 TATAAGTTGAACGACCTGAAGATGACGCCCC-CTGTGAGCAGTCACTATCTGCTCTTTGA 1036
Db 962 TATAAGTTGAACGACCTGAAGATGACGCCCCCTGTGAGCAGTCACTATCTGCTCTTTGA 903
QY 1037 TCAGATGACAGAGTTCATGACGAGAACCTGTGTGTATTACAGTACCAGGAGACACTTG 1096
Db 902 TCAGATGACAGAGTTCATGACGAGAACCTGTGTGTATTACAGTACCAGGAGACACTTG 843
QY 1097 GGGCTCTCGGATGACACTTCCAGCCGACGCTGAAGCAGTTCAGTCTTTTAAGTGAC 1156
Db 842 GGGCTCTCGGATGACACTTCCAGCCGACGCTGAAGCAGTTCAGTCTTTTAAGTGAC 783
QY 1157 CACACTCCAGAGGAGCTGTATGACTTTGCTAAGGAAATATAATGATGATGATGAGGG 1216
Db 782 CACACTCCAGAGGAGCTGTATGACTTTGCTAAGGAAATATAATGATGATGATGAGGG 723
QY 1217 AGAAGTTGTGGAATATGTGATGACCTTGTGAACCTGGAGGAGACGACGTAGCCACAGC 1276
Db 722 AGAAGTTGTGGAATATGTGATGACCTTGTGAACCTGGAGGAGACGACGTAGCCACAGC 663
QY 1277 AACCAAGAGACTTCTTTGGCTTGCAGAAACACAGATTTCTTTGCTTTTCCCAACA 1336
Db 662 AACCAAGAGACTTCTTTGGCTTGCAGAAACACAGATTTCTTTGCTTTTCCCAACA 603
QY 1337 GCCAGGCTTGTGATACCTCAGAGCTTCTCTTACTTCCAAAGTGAAGGAGAGCCCC 1396
Db 602 GCCAGGCTTGTGATACCTCAGAGCTTCTCTTACTTCCAAAGTGAAGGAGAGCCCC 543
QY 1397 CGTCTCTTAACCTGATGTCACTCAGGGGTGAGCCTGCTTTTCTATCTTCAACCTGCA 1456
Db 542 CGTCTCTTAACCTGATGTCACTCAGGGGTGAGCCTGCTTTTCTATCTTCAACCTGCA 483
QY 1457 CTTATGTTTCAACCTATCTTTCTACC-TTTTGTGATGAGTGTGCTCTCTTGC 1515
Db 482 CTTATGTTTCAACCTATCTTTCTACC-TTTTGTGATGAGTGTGCTCTCTTGC 423
QY 1516 CCAGGCTGGA-GTGCAATGSCAGTTCTCAGCTCACTGCAACCTCGCCTCTTGGTTCA 1574
Db 422 CCAGGCTGGAATGTCAGTGTCTCAGCTCACTGCAACCTCGCCTCTTGGTTCA 363
QY 1575 AGCAATCTGCTGATCAGCTCCGAGTACCTGGGATACAGGATGTGCCACACGCG 1634
Db 362 AGCAATCTGCTGATCAGCTCCGAGTACCTGGGATACAGGATGTGCCACACGCG 303
QY 1635 CGGCTAATTTTGTATTTTGTAGAGACGGGTTTGGCCATGTGGCCAGGCTGTCTCG 1694
Db 302 CGGCTAATTTTGTATTTTGTAGAGACGGGTTTGGCCATGTGGCCAGGCTGTCTCG 243
QY 1695 AACTCTTGACTTCAGATGATCCATCTGCTTGGCTTCCACAGTGTGGGATATACAGGG 1754
Db 242 AACTCTTGACTTCAGATGATCCATCTGCTTGGCTTCCACAGTGTGGGATATACAGGG 183
QY 1755 TGAGCCACCATGCCCCGGCTTTTCTACCTTTTACCTGTCTTTTATCCTCAGATCTG 1814
Db 182 TGAGCCACCATGCCCCGGCTTTTCTACCTTTTACCTGTCTTTTATCCTCAGATCTG 123
QY 1815 TTTTACACTTTCATCCCTGTCTCTCTC-ATGTTACACTTGTCTTCCCAATGTTCAATG 1873
```

```
Db 122 TTTTACACCTTCACTCCCTGTCTTCTCTATGTTCACTTGTCTCCCATGTTCTATAG 63
QY 1874 CTGCTCTTCTTACCATTTTGGTTTGAAGGAGCTTCTCTGCTTGTCTTTTGTGTTTTT 1933
Db 62 CTGCTCTTCTTACCATTTTGGTTTGAAGGAGCTTCTTCTGCTTGTCTTTTGTGTTTT 3
QY 1934 CC 1935
Db 2 CC 1

RESULT 3
AL569764/c
LOCUS
DEFINITION AL569764 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CSODI002YB17 3-PRIME, mRNA sequence.
ACCESSION AL569764
VERSION AL569764.2 GI:31291196
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li W.B., Gruber C., Jessee J. and Polayes D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT On Feb 16, 2001 this sequence version replaced gi:12925427.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqs@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by life technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 2233.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSODI002CH09NP1&cluster=2233.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSODI002CH09NP1.
FEATURES
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI002YB17"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 344 a 242 c 309 g 281 t 25 others
ORIGIN
Query Match 26.0%; Score 1004.8; DB 9; Length 1201;
Best Local Similarity 97.5%; Pred. No. 2.4e-14;
Matches 1021; Conservative 7; Mismatches 18; Indels 1; Gaps 1;
QY 922 GCTATCCGGTTGAGAAATTTGTGGCTACCATGATATCACTTGTGAGTTCGCTTATATA 981
Db 1046 GCTATCCAGTGAAGAA-TTGTGGCTACCATGATATCACTTGTGAGTTCGCTTATATA 998
QY 982 AGTTGAACGACCTTGAAGAAATGAGCCCTTGTGAGTTCGCTTATATA 1041
Db 987 AGTTGAACGACCTTGAAGAAATGAGCCCTTGTGAGTTCGCTTATATA 928
QY 1042 ATGACAAAGTTCATGACAGCAACCTGTGTATTACAGTACCAGGACACTTGGGGCC 1101
Db 927 ATGACAAAGTTCATGACAGCAACCTGTGTATTACAGTACCAGGACACTTGGGGCC 868
QY 1102 TCTCGGATGAGCACTTCCAGCCCGACGCTGAAGCAGTTCAGTCTTTAATGTGACCAC 1161
```



```
Db 867 TCTCGATGACACTTCAGCCGAGACTGAGCAGTTCAGTCTTTAATGTGACACAC 808
QY 1162 TCCAGAGAGAGTGTATGACTTTCAGGAAATATAATGATGATGATGAGGAGAAG 1221
Db 807 TCCAGAGAGAGTGTATGACTTTCAGGAAATATAATGATGATGATGAGGAGAAG 748
QY 1222 TTGTGGAATATGTGATGACCTCTTGGAACTGGAGGAGACAGCTAGCCACAGCAACCA 1281
Db 747 TTGTGGAATATGTGATGACCTCTTGGAACTGGAGGAGACAGCTAGCCACAGCAACCA 698
QY 1282 AAGAGACTTCTTCTTGGCGTTTCCAGAAACAGAGATCTTGTCTTTTCCACAGCCCA 1341
Db 697 AAGAGACTTCTTCTTGGCGTTTCCAGAAACAGAGATCTTGTCTTTTCCACAGCCCA 628
QY 1342 GGCTGTTGATACCTCAGAGCCTTCTCTTACTCTCCAAAGTGAAGGAGAGCCCGCTCT 1401
Db 627 GGCTGTTGATACCTAAGAGCCTTCTCTTACTCTCCAAAGTGAAGGAGAGCCCGCTCT 568
QY 1402 CTCTAACTGCATGTCATCAGGGGTGAGCTCGCTTTCCTATCTTCAACCTGCCACCTCA 1461
Db 567 CTCTAACTGCATGTCATCAGGGGTGAGCTCGCTTTCCTATCTTCAACCTGCCACCTCA 508
QY 1462 TGTTCACACCTATCTTCTCAGCTTCTTGTGAGATGAGTCTCGCTCTTTCGCCAGGC 1521
Db 507 TGTTCACACCTATCTTCTCAGCTTCTTGTGAGATGAGTCTCGCTCTTTCGCCAGGC 448
QY 1522 TGGAGTGAATAGGACAGTCTTCTCAGCTCACTCAACCTCGCTCTTGGGTTCAAGCAATT 1581
Db 447 TGGAGTGAATAGGACAGTCTTCTCAGCTCACTCAACCTCGCTCTTGGGTTCAAGCAATT 388
QY 1582 CTGCTGCATCAGCTCCCGAGTACTGGGATTCAGGATGTCACACAGCCCGGCTAA 1641
Db 387 CTGCTGCATCAGCTCCCGAGTACTGGGATTCAGGATGTCACACAGCCCGGCTAA 328
QY 1642 TTTTGTATTTTAGTAGAGAGCGGGTTTGGCATGTTGGCAGGCTGTCTCGAATCTTT 1701
Db 327 TTTTGTATTTTAGTAGAGAGCGGGTTTGGCATGTTGGCAGGCTGTCTCGAATCTTT 268
QY 1702 GACTTCAGATGATCATCTGCTTGGCTTCCAGTCTCCAGTCTGGGATTCAGGCGTGAGCCA 1761
Db 267 GACTTCAGATGATCATCTGCTTGGCTTCCAGTCTCCAGTCTGGGATTCAGGCGTGAGCCA 208
QY 1762 CCATGCCCGGCTCTTCTCAGCTTACACCTGCTCTTATCTCATCTCATCTGTTTTCAC 1821
Db 207 CCATGCCCGGCTCTTCTCAGCTTACACCTGCTCTTATCTCATCTCATCTGTTTTCAC 148
QY 1822 ACCTTCATCCCTGCTTCTCCTCATGTTTCACTTGTCTTCCCATGTTTTCATAGTGCCTTT 1881
Db 147 ACCTTCATCCCTGCTTCTCCTCATGTTTCACTTGTCTTCCCATGTTTTCATAGTGCCTTT 88
QY 1882 CTTACCATTTTGGTTGAGGCGAGTCTTCTCTGCTGTTTGTGTTTGTGTTTTCACGAAA 1941
Db 87 CTTACCATTTTGGTTGAGGCGAGTCTTCTCTGCTGTTTGTGTTTGTGTTTTCACGAAA 28
QY 1942 ATCAGTATTTATTTTAAATGAAGAAA 1968
Db 27 ATAAATTATWNTWNTAAATAATAAAAA 1
```

```
RESULT 4
AL517454/c 1196 bp mRNA linear EST 12-MAY-2003
LOCUS
DEFINITION AL517454 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
CSODA003YF11 3-PRIME, mRNA sequence.
ACCESSION AL517454
VERSION AL517454.2 GI:30534773
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1196)
```

AUTHORS
TITLE
JOURNAL
COMMENT

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 13, 2001 this sequence version replaced gi:12780947.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 2233.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cmap.cgi?seq=CSODA003CC06NP1&cluster=2233.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSODA003CC06NP1.

FEATURES

source

1..1196
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODA003YF11"
/tissue_type="NEUROBLASTOMA"
/clone_lib="Homo sapiens NEUROBLASTOMA"
/note="Vector: PCWSPORT_6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the PCWSPORT 6 vector.
Library was not normalized."
BASE COUNT 339 a 246 c 312 g 264 t 35 others
ORIGIN

Query Match 25.5%; Score 985.2; DB 9; Length 1196;
Best Local Similarity 97.2%; Pred. No. 6.4e-14;
Matches 1011; Conservative 8; Mismatches 19; Indels 2; Gaps 2;
QY 885 ATACAGTGTGAAGAGAACCTCACCCAGTTATAGGAGCTATCCCGTTGAGAAATTTGTG 944
Db 1038 AAATACGWTGTTGAKAKACCTCACCCAGTTATAGAGCTATCCCGTTGAGAAATTTKT- 980
QY 945 GTACACATGATATCATTTACTTTGAGTTTGCCTTATTAAGTTGAACGACTGAAGATGCA 1004
Db 979 GCTACCATGATATCATTTACTTTGAGTTTGCCTTATTAAGTTGAACGACTGAAGATGCA 920
QY 1005 GCCCCTGTGAGTACAGTATCTGCTCTTTGATCAGATGACAGGTTCATGACAGAAC 1064
Db 919 GCCCCTGTGAGTACAGTATCTGCTCTTTGATCAGATGACAGGTTCATGACAGAAC 860
QY 1065 CTGGTGTATTTACAGTACACAGGAGACACTTGGGCTCTCGGATGAGCCTTCCAGCCC 1124
Db 859 CTGGTGTATTTACAGTACACAGGAGACACTTGGGCTCTCGGATGAGCCTTCCAGCCC 801
QY 1125 AGACCTGAAGCAGTTTCAGTTCTTTAATGTGACCACTCCAGAGGAGCTGTATGACCTT 1184
Db 800 AGACCTGAAGCAGTTTCAGTTCTTTAATGTGACCACTCCAGAGGAGCTGTATGACCTT 741
QY 1185 GCTAAGGAAATATTAATGATGATGAGGAGAGTTTGGAAATATGTGATGACCTC 1244
Db 740 GCTAAGGAAATATTAATGATGATGAGGAGAGTTTGGAAATATGTGATGACCTC 681
QY 1245 TTGGAATGAGGAGACAGCTAGCCACAGCAACCAAGAGACTTCTCTTGGGTTCA 1304
Db 680 TTGGAATGAGGAGACAGCTAGCCACAGCAACCAAGAGACTTCTCTTGGGTTCA 621
QY 1305 GGAACACAGATTTCTTTGCTTTTCCCAACAGCCAGGCTGTTGATACCTCAGAGCCTT 1364
Db 620 GGAACACAGATTTCTTTGCTTTTCCCAACAGCCAGGCTGTTGATACCTCAGAGCCTT 561
QY 1365 CTCTTTACTCTCCAAAGTGAAGGAGAGCCCGCTCTCTTAACGTGATGTCATGAGGG 1424
Db 560 CTCTTTACTCTCCAAAGTGAAGGAGAGCCCGCTCTCTTAACGTGATGTCATGAGGG 501
QY 1425 TGAGCCTGCTTCTCTATCTTTCACACCTGCGCACTCATGTTTCCACACCTATCTTTCTCACC 1484

```
Db 500 TGAGCGTGCCTTCTCTATCTTACACACCTGCGACCTCATGTTTACACACCTATCTTCTCACC 441
Qy 1485 TTTTCTTTTGTAGATGAGTCTGCTCTCTTGTCCAGGCTGGAGTGCATAGGACGCTTCTCA 1544
Db 440 TTTTCTTTTGTAGATGAGTCTGCTCTCTTGTCCAGGCTGGAGTGCATAGGACGCTTCTCA 381
Qy 1545 GTCCTACTGCAACCTCCGCTCTTGGTTCACGCAATCTGTGTCATCAGCCTCCCGAGTA 1604
Db 380 GTCCTACTGCAACCTCCGCTCTTGGTTCACGCAATCTGTGTCATCAGCCTCCCGAGTA 321
Qy 1605 CTGGGATTTACAGGATGTCACACACGCGCGGCTAAATTTTGTATTTTGTAGAGAGCG 1664
Db 320 CTGGGATTTACAGGATGTCACACACGCGCGGCTAAATTTTGTATTTTGTAGAGAGCG 261
Qy 1665 GGTTTTGGCATGTTGGCCAGGCTGCTCGAACTCTTGACTTCAGATCCATCTGCGCT 1724
Db 260 GGTTTTGGCATGTTGGCCAGGCTGCTCGAACTCTTGACTTCAGATCCATCTGCGCT 201
Qy 1725 TGGCTTCCACAGTGTGGGATTAAGGCGTGAGCAATGCCCGGCTCTTCTCACC 1784
Db 200 TGGCTTCCACAGTGTGGGATTAAGGCGTGAGCAATGCCCGGCTCTTCTCACC 141
Qy 1785 TTTTACACCTGCTCTTATTCCTCACATCTGTTTTCACACCTTCATCCCTGCTCTCTCAT 1844
Db 140 TTTTACACCTGCTCTTATTCCTCACATCTGTTTTCACACCTTCATCCCTGCTCTCTCAT 81
Qy 1845 GTTTCACACTGTCTTCCCATGTTTCATAGCTGCTCTTCTTACATTTTGGTTTGAAGGCG 1904
Db 80 GTTTCACACTGTCTTCCCATGTTTCATAGCTGCTCTTCTTACATTTTGGTTTGAAGGCG 21
Qy 1905 AGTCTTCTGCTGCTGTTTT 1924
Db 20 AGTCTTCTGCTGCTGTTTT 1
```

```
RESULT 5
AL514488 1201 bp mRNA linear EST 08-MAY-2003
LOCUS AL514488 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
DEFINITION CLOBB003ZF09 5-PRIME, mRNA sequence.
ACCESSION AL514488
VERSION AL514488.2 GI:30464373
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 13, 2001 this sequence version replaced gi:12777982.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 2233.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CLOBB003ZF09RP1&cluster=2233.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue Genoscope sequence ID : CLOBB003ZF09RP1.
Location/Qualifiers
1. 1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CLOBB003ZF09"
/tissue_type="NEUROBLASTOMA"
/clone_lib="Homo sapiens NEUROBLASTOMA"
```

FEATURES

source

/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

BASE COUNT 274 a 299 c 324 g 246 t 58 others
ORIGIN

Query Match 25.3%; Score 979.6; DB 9; Length 1201;
Best Local Similarity 94.0%; Pred. No. 8.4e-14;
Matches 1025; Conservative 31; Mismatches 31; Indels 4; Gaps 4;

Qy 59 CGCGATGAGAGCCGCGGCGCGCGCGCGCGCGCTGTAGCGCTCTCTGTGGTGGCGCTG 118
Db 58 CGCGATGAGAGCCGCGGCGCGCGCGCGCGCGCTGTAGCGCTCTCTGTGGTGGCGCTG 117
Qy 119 CGCGCTGCGCGCGCGCGCGCGCGCGCGCGCTGTAGCGCTCTCTGTGGTGGCGCTG 178
Db 118 CGCGCTGCGCGCGCGCGCGCGCGCGCGCGCTGTAGCGCTCTCTGTGGTGGCGCTG 177
Qy 179 CGAGCTGATCCGCTCGAGTCGCGCTTACCGGCGCGCGCTGTAGCGCTCTCTGTGGTGGCGCTG 238
Db 178 CGAGCTGATCCGCTCGAGTCGCGCTTACCGGCGCGCGCTGTAGCGCTCTCTGTGGTGGCGCTG 237
Qy 239 CTGGCGCGAGAGCTGCGCTTACCGCTGAGATGAGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 298
Db 238 CTGGCGCGAGAGCTGCGCTTACCGCTGAGATGAGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 297
Qy 299 CAGCGAGCGCTTCTGCGCGCGCGCGCGCGCGCGCGCGCTGCGCTGCGCTGCGCTGCGCTG 358
Db 298 HAGCGAGCGCTTCTGCTTACCGCACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 357
Qy 359 CTTGCGCGAGCTATCCGAGCTGCGCTTCTGCGCGCGCGCTGCTGCGCGCGCGCGCTGCTG 418
Db 358 CTTGCGCGAGCTATTCGAGCTGCGCTTCTGCGCGCGCGCTGCTGCGCGCGCGCGCGCTGCTG 416
Qy 419 CAAGCGCTGCAAGCAGCGCGCTGCGCGCTTCCGCGCGCTGCGCGCGCGCGCGCGCGCGCTGCT 478
Db 417 CAAGCGCTGCAAG-AGGCGCTTGCWAGCTTTCGAGTTCWAGCTGCGCGCGCGCGCTGCT 475
Qy 479 GCGGAGCTTCCAGCGCGCGCGCGCGCTTACAGCTTCTGCGCTGCTTCTGCTTCTGCTTCTG 538
Db 476 GCGGAGCTTTHAGCGCGCGCGCGCGCTTAAATTTCTGCGCTGCTTCTGCTTCTGCTTCTG 535
Qy 539 TAATCTCCCAAGCGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 598
Db 536 TAATATMCCCAAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 595
Qy 599 AATGATGAAGAGGAAACATGCTATTTATTAAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 658
Db 596 AATGATGAAGAGGAAACATGCTATTTATTAARARCATGCTGCTGCTGCTGCTGCTGCTGCT 655
Qy 659 AGACTCGGAACCAAGTCATATGAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 718
Db 656 AGACTCGGAACCAAGTCATATGAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 715
Qy 719 TGAGAACTGGAGAACTCCATCATCAGACATGAGAGCTGGCGCTTCCGCGCTTCTTCTTCAAGC 778
Db 716 TGAGAACTGGAGAACTCCATCATCAGACATGAGAGCTGGCGCTTCCGCGCTTCTTCTTCAAGC 775
Qy 779 CTTTTCAGAGTGTCTCCAGCGCTGCGAGGCTTCCAGGAGATCAAGGAGCTTCAAGGATTT 838
Db 776 CTTTTCAGAGTGTCTCCAGCGCTGCGAGGCTTCCAGGAGATCAAGGAGCTTCAAGGATTT 835
Qy 839 CTACCTTTCATAGCAGATCATTTATGTAGAAGTCTTCTGGAATGCAAAATACAGTGTGAAGA 898
Db 836 CTACCTTTCATAGCAGATCATTTATGTAGAAGTCTTCTGGAATGCAAAATACAGTGTGAAGA 895
Qy 899 GAACCTCACCCCGAGTTATAGAGGCTATCCGGTTTGAAGAAATTTGTGGCTACCATGTATCA 958
Db 896 GAACCTCACCCCGAGTTATAGAGGCTATCCGGTTTGAAGAAATTTGTGGCTACCATGTATCA 955
Qy 959 TTACTTGCAGTTTGCCTTATTATTAAGTTGAACGAGCTTGAAGAAATGACAGCCCGCTGTCAGT 1018

```

1133 TCAGTTCTTTA 1149
1134 YCAGTTTAA 1144

AL556976 Homo sapiens 1201 bp mRNA linear EST 31-MAY-2003
AL556976 Homo sapiens T CELLS (JURKAT CELL LINE) Homo sapiens cDNA
LOCUS CS0DH005Y101 5-PRIME, mRNA sequence.
DEFINITION
ACCESSION AL556976
VERSION 1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 15, 2001 this sequence version replaced gi:12900137.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
2233.f For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DH005AE01QP1&cluster=2233.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DH005AE01QP1.
FEATURES
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DH005Y101"
/tissue_type="T CELLS (JURKAT CELL LINE)"
/cell_line="JURKAT CELL LINE"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE)"
/note="vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
BASE COUNT 272 a 339 c 320 g 234 t 36 others
ORIGIN
Query Match 25.1%; Score 969.8; DB 9; Length 1201;
Best Local Similarity 97.7%; Pred. No. 1.3e-13;
Matches 1015; Conservative 9; Mismatches 11; Indels 4; Gaps 4;

65 GGAGCCGGGGCGCGGGGGCGCGGGCGCTGGTGGTGGCTGGCGGCT 124
58 GGATCCGGGGCGCGGGGGCGCGGGCGCTGGTGGTGGCTGGCGGCT 117
125 GCGGGCGGGGGCGCGGGCGCGGGCGCTGGTGGTGGCTGGCGGCT 184
118 GC-GCGGGGGCGG-CCAATACGAACGCTACAGCTTCCGAGCTTCCACGGACGAGCT 175

```

```

185 GATCGCGCTCAGTTCGGCTTACCGGCAACGCGCTTGGCAAGTACAGCGGGAGACCTGGGC 244
176 GATCGCGCTCAGTTCGGCTTACCGGCAACGCGCTTGGCAAGTACAGCGGGAGACCTGGGC 235
245 CGAGAGCGTGGCTTACCTGGAGATCAGCTTCGGCTTGGCAAGTACAGCGGGAGACCTGGGC 304
236 CGAGAGCGTGGCTTACCTGGAGATCAGCTTCGGCTTGGCAAGTACAGCGGGAGACCTGGGC 234
305 GGCCTTCTGCGCACCGCAATGTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 364
295 GGCCTTCTGCGCACCGCAATGTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 354
365 CAGCTATCCGAGCTGGCGCTTTCGGGGCGCTTTCGGGGCGCTTTCGGGGCGCGCGCGCGCG 424
355 CAGCTATCCGAGCTGGCGCTTTCGGGGCGCTTTCGGGGCGCTTTCGGGGCGCGCGCGCGCG 414
425 CTGCAAGCAGCGGGCTGCGCAGCTTTCGGCAGTCCCGCAGTCCCGCAGCGCGCGCGCGCG 484
415 CTGCAAGCAGCGGGCTGCGCAGCTTTCGGCAGTCCCGCAGTCCCGCAGCGCGCGCGCGCG 474
485 CTTCCAGCGCGCGGAGCGCTTCAAGTTCTTCCAGTTGGCTTCTTCAAGCAAAATATCT 544
475 CTTCCAGCGCGGAGCGCTTCAAGTTCTTCCAGTTGGCTTCTTCAAGCAAAATATCT 534
545 CCCCAAGCAGCGGGCTGCGCAGCTTTCGGCAGTCCCGCAGTCCCGCAGCGCGCGCGCG 604
535 CCCCAAGCAGCGGGCTGCGCAGCTTTCGGCAGTCCCGCAGTCCCGCAGCGCGCGCGCG 594
605 GAAGAGAAACATGATATTAAGAGCTTCTTCAAGCTTCTTCAAGCAAAATATCT 664
595 GAAGAGAAACATGATATTAAGAGCTTCTTCAAGCTTCTTCAAGCAAAATATCT 654
665 GGAACCAAGTCAATGAAAGCTTTCATCCGAGCAGTTCGGGCGCATACAAAGTGAAGAA 724
655 GGAACCAAGTCAATGAAAGCTTTCATCCGAGCAGTTCGGGCGCATACAAAGTGAAGAA 714
725 CTGAGAGCAATCCATCAAGAGCTTCTTCAAGCTTCTTCAAGCAAAATATCT 784
715 CTGAGAGCAATCCATCAAGAGCTTCTTCAAGCTTCTTCAAGCAAAATATCT 774
785 CGAGTGTCTCGACCGCTTCGAGGGTTCAGGAGATCAAGGAGTTCAGGATTTTACCT 844
775 CGAGTGTCTCGACCGCTTCGAGGGTTCAGGAGATCAAGGAGTTCAGGATTTTACCT 834
845 TTCCATAGCAGATCATTTATGTAAGTTCTGGAATGCAAAATACAGTGTGAAGAGACCT 904
835 TTCCATAGCAGATCATTTATGTAAGTTCTGGAATGCAAAATACAGTGTGAAGAGACCT 894
905 CACCCAGTTATAGAGCTATCCGGTTGAGAAATTTGGCTACCATGTATCATTTACTT 964
895 CACCCAGTTATAGAGCTATCCGGTTGAGAAATTTGGCTACCATGTATCATTTACTT 954
965 GCAGTTTCTCTTATTAAGTTGAACGACCTGAAGATGCAAGCGCCCTGTGAGTCAAGTCA 1024
955 GCAGTTTCTCTTATTAAGTTGAACGACCTGAAGATGCAAGCGCCCTGTGAGTCAAGTCA 1014
1025 TCTGCTCTTTTATGAGAAATGCAAGGCTCATGAGCAAGCACTGTGTGATTAACGATACCA 1084
1015 TCTGCTCTTTTATGAGAAATGCAAGGCTCATGAGCAAGCACTGTGTGATTAACGATACCA 1073
1085 CAGGGAGACACTTGGGGCGCTC 1103
1074 MAGGACACTTGGGGCYTC 1092

RESULT 7
BX425905 Homo sapiens 1201 bp mRNA linear EST 15-MAY-2003
LOCUS CS0DA004Y407 5-PRIME, mRNA sequence.
DEFINITION
ACCESSION BX425905
VERSION 1
GI:30774508

```



```
FEATURES
source      Location/Qualifiers
1.1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSDA004Y07"
/tissue_type="NEUROBLASTOMA"
/clone_lib="Homo sapiens NEUROBLASTOMA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dt) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
BASE COUNT 309 a 234 c 320 g 275 t 63 others
ORIGIN
Query Match 24.7%; Score 954.8; DB 13; Length 1201;
Best Local Similarity 90.6%; Pred. No. 2.8e-13;
Matches 1006; Conservative 27; Mismatches 74; Indels 3; Gaps 2;
QY 824 GGACTTCAAGGATTTCTACCTTCCATAGCAGATCATATGTAGAAGTTCGGAATGCAA 883
DB 1110 GGGGTCAGGGRAYAGRCCTYAAAGRTTYCCCTTCCAWRSARAYCATWTKARPAATYKG 1051
QY 884 AATACAGTGTGACGAGAACCTCACCCAGATTATAGGAGCTATCCGGTTGAGAAATTTGT 943
DB 1050 AATGCAAAACARKKGAARAAACCCCCCAKTAWRGAGCTACCCGGTGAGAAATYKG 991
QY 944 GGCTACCATGATCATTAATCTTGGCTTGGCTTATATAGTTGAAGCACTGAAGAAATGC 1003
DB 990 GGCTACCATGATCATTAAT--GCAKTTGGCTTATATAG--TGAACGACCTGAAGAAATGC 934
QY 1004 AGCCCGCTGTCAGTACGATCTGCTTGTGATGCAAGTACAAAGCTATGACAGAGAA 1063
DB 933 AGCCCGCTGTCAGTACGATCTGCTTGTGATGCAAGTACAAAGCTATGACAGAGAA 874
QY 1064 CCTGGTGTATTAACGATACCAAGGACACTTGGGGCTCTCGGATGAGCACTTCAGACC 1123
DB 873 CCTGGTGTATTAACGATACCAAGGACACTTGGGGCTCTCGGATGAGCACTTCAGACC 814
QY 1124 CAGACCTGAACAGTTCAGTCTTTTATGTGACCACTCAGAGAGAGCTGTATGACCT 1183
DB 813 CAGACCTGAACAGTTCAGTCTTTTATGTGACCACTCAGAGAGAGCTGTATGACCT 754
QY 1184 TGCTAAGGAAATATATGATGATGATGAGGAGAGAGTGTGGAATATGCGATGACCT 1243
DB 753 TGCTAAGGAAATATATGATGATGATGAGGAGAGAGTGTGGAATATGCGATGACCT 694
QY 1244 CTTGGAAGTGAAGGAGACCACTAGCCACAGCAACCAAGAGACTTCCTCTTGGGGTTC 1303
DB 693 CTTGGAAGTGAAGGAGACCACTAGCCACAGCAACCAAGAGACTTCCTCTTGGGGTTC 634
QY 1304 AGGAACACAGATCTTTGCTCTTTTCCACAGCCAGGCTGTGATACCTCAGAGCCT 1363
DB 633 AGGAGACACAGATCTTTGCTCTTTTCCACAGCCAGGCTGTGATACCTCAGAGCCT 574
QY 1364 TCTCTTTACTCCAAAGTGAAGGAGAGCCCGCTCTCTCTAACTGCATGTCAATCAGG 1423
DB 573 TCTCTTTACTCCAAAGTGAAGGAGAGCCCGCTCTCTCTAACTGCATGTCAATCAGG 514
QY 1424 GTGAGCCTGCCTTCTTCTATCTTCACTGCACTCATGTGTCACACTATCTTCTCAC 1483
DB 513 GTGAGCCTGCCTTCTTCTATCTTCACTGCACTCATGTGTCACACTATCTTCTCAC 454
QY 1484 CTTTTTTTTGAGATGGAGTCTCGTCTCTCTTCCAGGCTGGAGTGCATGSCAGCTTCTC 1543
DB 453 CTTTTTTTTGAGATGGAGTCTCGTCTCTCTTCCAGGCTGGAGTGCATGSCAGCTTCTC 394
QY 1544 AGCTCACTGCACACTCCCGCTCTTGGGTTCAAGCAATCTCTGTCATCAGCTCCAGT 1603
DB 393 AGCTCACTGCACACTCCCGCTCTTGGGTTCAAGCAATCTCTGTCATCAGCTCCAGT 334
QY 1604 ACCTGGGATTACAGGCAATGTGCCACACGCGCGGCTAATTTTGTATTTTAGTAGAGAG 1663
```

```

333 ACCTGGGATWACAGGCATGTGCCACACGCGCGCTAATTTGTATTTAGTAGAGCG 274
1664 GGGTTTTGCCATGTGTGCCAGGCTGGTCTCGAACTCTTGACTTCAGATGATCCATCTGCC 1723
273 GGGTTTTGCCATGTGTGCCAGGCTGGTCTCGAACTCTTGACTTCAGATGATCCATCTGCC 214
1724 TTGGCCTCCACAGTGTGGGATACAGGGGTGAGCCACCATGCCCGGCTCTTCTTCAC 1783
213 TTGGCCTCCACAGTGTGGGATACAGGGGTGAGCCACCATGCCCGGCTCTTCTTCAC 154
1784 CTTTACACCTGTCTTCTTATCTCCTCACATCTGTGTTTTCACACTTCATCCCTGCTCTCA 1843
153 CTTTACACCTGTCTTCTTATCTCCTCACATCTGTGTTTTCACACTTCATCCCTGCTCTCA 94
1844 TGTTCACACTTGTCTTCCCATGTTTCATAGCTGCTTCTTACCATTTGGTTTGAAGGG 1903
93 TGTTCACACTTGTCTTCCCATGTTTCATAGCTGCTTCTTACCATTTGGTTTGAAGGG 34
1904 CAGTCTTCTGCTGCTGTTTGTGTTTGTGTTT 1933
33 CAGTCTTCTGCTGCTGTTTGTGTTTGTGTTT 4
RESULT 9
BU902192 1148 bp mRNA linear EST 17-OCT-2002
LOCUS AGENCOURT_10127451 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6502891
DEFINITION 5', mRNA sequence.
ACCESSION BU902192
VERSION BU902192.1 GI:24084105
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1148)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14059 row: k column: 20
High quality sequence stop: 754.
Location/Qualifiers
1.1148
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6502891"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 71"
/notes="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.1 kb"
BASE COUNT 253 a 363 c 315 g 216 t 1 others
ORIGIN
Query Match 24.6%; Score 954; DB 13; Length 1148;
Best Local Similarity 96.2%; Pred. No. 3e-13;
Matches 1008; Conservative 2; Mismatches 32; Indels 6; Gaps 3;
QY 67 AGCCGGGGCGCGGGGGCGCGCGCTGCTAGCCCTGCTGCGCTGCGCTGCGCTGCGCTGCG 126
DB 1 AGCCGGGGCGCGGGGGCGCGCGCTGCTAGCCCTGCTGCGCTGCGCTGCGCTGCGCTGCG 60
```

```
QY 127 GCGCGGCGCGCGCCCAATACGAACCTACAGCTTCGCGAGCTTCCACGGACGAGCTGA 186
Db 61 GCGCGGCGCGCGCCCAATACGAACCTACAGCTTCGCGAGCTTCCACGGACGAGCTGA 120
QY 187 TCGCGCTGAGTCGCGCTACCGGACGCGCTGGACAGTACAGCGGGGAGCACTGGGCGG 246
Db 121 TCGCGCTGAGTCGCGCTACCGGACGCGCTGGACAGTACAGCGGGGAGCACTGGGCGG 180
QY 247 AGAGCTGCGCTACCTGAGATCAGCTTCGCGCTGCGAGCTTCGCGCGGACGAGG 306
Db 181 AGAGCTGCGCTACCTGAGATCAGCTTCGCGCTGCGAGCTTCGCGCGGACGAGG 240
QY 307 CTTCTGCGACGCGAATCGAGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCTCGCCA 366
Db 241 CTTCTGCGACGCGAATCGAGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCTCGCCA 300
QY 367 GCTATCCCGAGCTGCGCTCTTCGCGGCGCTGCTGCGCGCGCGCGAGCTTCGCGGCT 426
Db 301 GCTATCCCGAGCTGCGCTCTTCGCGGCGCTGCTGCGCGCGCGCGAGCTTCGCGGCT 360
QY 427 GCAAGCAGCGCGCTGCGAGCTTCGCGCGAGTCCGCGAGTCCGCGAGCGCGAGCTGCGGACT 486
Db 361 GCAAGCAGCGCGCTGCGAGCTTCGCGCGAGTCCGCGAGTCCGCGAGCGCGAGCTGCGGACT 420
QY 487 TCCAGCGCGCGAGCGCGCTACAGCTTCGCGAGTTCGCGAGTTCGCGAGCGCGAGTTCGCG 546
Db 421 TCCAGCGCGCGAGCGCGCTACAGCTTCGCGAGTTCGCGAGTTCGCGAGCGCGAGTTCGCG 480
QY 547 CAAAGCAGCTGCGCGCTGCTCACACTTCTACTGAGCATCTCTGATGACGAAATGATGA 606
Db 481 CAAAGCAGCTGCGCGCTGCTCACACTTCTACTGAGCATCTCTGATGACGAAATGATGA 540
QY 607 AGAGAACATGCGATATATAGAGCGCTGCTGCGCGAGCTACATTAAGACCTGG 666
Db 541 AGAGAACATGCGATATATAGAGCGCTGCTGCGCGAGCTACATTAAGACCTGG 600
QY 667 AAACCAAGTCATATGAAGCGCTTTCATCCGAGCAGTGGCGGATACAGCGTGAGAACT 726
Db 601 AAACCAAGTCATATGAAGCGCTTTCATCCGAGCAGTGGCGGATACAGCGTGAGAACT 660
QY 727 GGAGAACATCCATCAGACATGAGAGCTGGCGCTTCCGAGCTTCTTCAAAGCTTTTACG 786
Db 661 GGAGAACATCCATCAGACATGAGAGCTGGCGCTTCCGAGCTTCTTCAAAGCTTTTACG 720
QY 787 AGTGCTCGCGAGCTGCGAGGTTCCAGGAGATCAAGGACTTCAAGGATTTCTACCTTT 846
Db 721 AGTGCTCGCGAGCTGCGAGGTTCCAGGAGATCAAGGACTTCAAGGATTTCTACCTTT 780
QY 847 CCATAGCAGATCATTTATGTAAGTTCTGGAATGCAAAATACAGTGTGAAGAGAACCTCA 906
Db 781 CCATAGCAGATCATTTATGTAAGTTCTGGAATGCAAAATACAGTGTGAAGAGAACCTCA 840
QY 907 CCCAGTTATAGAGGCTATCCGTTGAGAAATTTGCGCTACCATGATCATTTACTTGC 966
Db 841 CCCAGTTATAGAGGCTATCCGTTGAGAAATTTGCGCTACCATGATCATTTACTTGC 900
QY 967 AGTTTGCCTATTATAGTTGAAGCAGCTGAAGAAATGCG - CCGCTTGTGAGTCAAGCTAT 1025
Db 901 AGTTTGCCTATTATAGTTGAAGCAGCTGAAGAAATGCGCCCGCCCTGTGAGGCGCAGTAT 960
QY 1026 CTGCTCTTTGATCAGAAATGCAAGGTCATGCAGC - AGAAGCTGGTGTATTAACGATACC 1083
Db 961 CGGCGCTTTGATCAGAAAGACAGGCTCTGCGACCAACAACTGGGGGTATTATCCCGCTCC 1020
QY 1084 ACAGGGACAC - - - TTGGGCGCTTCGGA 1108
Db 1021 CCCCGGGGACCTTTTGGGGCGCTTCGGA 1048
```

RESULT 10
BX380662/c
LOCUS BX380662 1046 bp mRNA linear EST 08-MAY-2003

DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BX380662 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0DI056Y117 3-PRIME, mRNA sequence.
BX380662
BX380662.1 GI:30443535
EST.
Homo sapiens (human)
Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1046)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 2233.f. For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DI056AE09NP1&cluster=2233.f. Contact :
Feng Liang Email : fliang@life.techn.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DI056AE09NP1.

FEATURES
Source

Location/Qualifiers

1..1046

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DI056Y117"

/tissue_type="PLACENTA COT 25-NORMALIZED"

/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/notes="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched double-strand cDNA was

digested with NotI and cloned into the NotI and EcoRV

sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 304 a 210 c 286 g 226 t 20 others

ORIGIN

Query Match 24.5% Score 947.2; DB 13; Length 1046;

Best Local Similarity 97.9%; Pred No 4.5e-13;

Matches 953; Conservative 10; Mismatches 9; Indels 1; Gaps 1;

QY 969 TTTGCCCTATTATAGTTGAAACGACCTGAGAAATGCGAGCCCTGTGCGAGTACGATCTG 1028

Db 980 KTTGCCCTTTTATAGTTGAAACGACCTGAGAAAGCGCCCTGTGCGAGTACGATCTG 921

QY 1029 CTCTTTGATCAGATGACAGAGTTCATGACAGCACTTCCAGCCAGACCTGAAAGCAGTTCAGTTCCTTT 1088

Db 920 CTCTTTGATCAGATGACAGAGTTCATGACAGCACTTCCAGCCAGACCTGAAAGCAGTTCAGTTCCTTT 861

QY 1089 GACACTTTGGGCGCTCTCGGATGAGCACTTCCAGCCAGACCTGAAAGCAGTTCAGTTCCTTT 1148

Db 860 RACACTTTGGGCGCTCTCGGATGAGCACTTCCAGCCAGACCTGAAAGCAGTTCAGTTCCTTT 801

QY 1149 AATGTGACCACTCCAGAGAGGAGCTGTATGCTTCTGAAGAAATATATAGTATGAT 1208

Db 800 AATGTGACCACTCCAGAGAGGAGCTGTATGCTTCTGAAGAAATATATAGTATGAT 741

QY 1209 GATCAGGAGAGAGTTGGAATATGCGATGACCTCTTGAAGCTGAGGAGAGACAGCTAG 1268

Db 740 GATCAGGAGAGAGTTGGAATATGCGATGACCTCTTGAAGCTGAGGAGAGACAGCTAG 681

QY 1269 CCCACAGCAACCAAGAGAGCTTCTCTTGGCGTTGAGAAACACAGATTCCTTTGTCCTTT 1328

Db 680 CCCACAGCAACCAAGAGAGCTTCTCTTGGCGTTGAGAAACACAGATTCCTTTGTCCTTT 621

QY 1329 TCCCAAGAGCCGAGCTTGTGATACCTCAGAGCTTCTCTTCTCCAAAGTGAAGG 1388

Db 620 TCCCAAGAGCCGAGCTTGTGATACCTCAGAGCTTCTCTTCTCCAAAGTGAAGG 561

QY 1389 GAAGCCCGCTCTCTCTAACTGATGTCATCAGGGGTGAGCCTGCTTCTCTATCTTCTAC 1448


```
Db 9 TTCTAGGCTGCGCGCTCGCGGTCTCTAGCTCGGGCTCTGTGTTGTTGGCCACGCGC 68
Qy 62 GATGAGCCGCGGGCGCGCGGGCGCGCGCTGCTAGCGCTGCTGTGGTGGCTGCGC 121
Db 69 GATGGGCCCCGAGCCCTACGCCC---GGTTGCTGGTCTGCTGTGTGGTGGC 125
Qy 122 GCTGCGCGCGGGCGCGCCCAATACGAACCTACAGCTTCGGAGCTTCCAGCTGCGGACGA 181
Db 126 GCGACACCCCGGGCGGGGAGTACGAGCGCTACAGCTTCGGAGCTTCCGCGGACGA 185
Qy 182 GCTGATCCGCTCGAGTCGCGCTACCGGCACGGCTGGACAGTACAGGGCGAGCACTG 241
Db 186 GCTGATCCGCTCGAGTCGCGCTACCGGCACGGCTGGACAGTACAGGGCGAGCACTG 245
Qy 242 GCGCGAGAGSTKGGCTACTGAGATCAGCTTGGGCTGCGAGCTGCTGGCGGACAG 301
Db 246 GCGCGAGAGCTGGCTACCTGAGGTGAGCTTGGGCTGCGGCTGCGGCTGCGGACAG 305
Qy 302 CGAGGCTTCTGCCACCGCAACTGACGCGCGCGCGCGAGCCGCGCGCGCGCT 361
Db 306 CGAGGCTTCTGCCACCGCAACTGACGCGCGCGCGCGAGCCGCGCGCGCGCGCT 365
Qy 362 CGCGAGTATCCGAGCTGCGCTTCTTGGGGGCTGCTGCGCGCGCGAGCTGCTCAA 421
Db 366 GCGCGAGCGCGGAGCTTCTTGGGAGGCTGCTGCGCGCGCGAGTGCCTCAA 425
Qy 422 GCGCTGCAAGCAGGCTGCGAGCTTCCGCGAGTCCGCGCGCGCGAGGCTGCTGC 481
Db 426 GCGCTGCAAGCAGGCTGCGAGCTTCCGCGAGTCCGCGCGCGAGGCTGCTGC 485
Qy 482 GGACTTCCAGCGCGCGAGCTTCAAGTTCCTGCGAGTTCGCTTACTTCAAGCAATAA 541
Db 486 GGACTTCCAGCGCGAGCTTCAAGTTCCTGCGAGTTCGCTTACTTCAAGCAATAA 545
Qy 542 TCTCCCAAGCCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 601
Db 546 CTTCCCAAGCCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 605
Qy 602 GATGAAGAGGAACATGCTATTAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 661
Db 606 GATGAAGAGGAACATGCTATTAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 665
Qy 662 CTTGGAACCAAGTCAATGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 721
Db 666 CTTGGAACCAAGTCAATGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 725
Qy 722 GACTGGAGACATCCATCAGACATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 781
Db 726 GACTGGAGACATCCATCAGACATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 785
Qy 782 TTACGAGTGTCTCGAGCTGCTGAGGCTTCCAGGAGATCAAGGACTTCAAGGATTCTA 841
Db 786 CTACGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 845
Qy 842 CTTTCCATAGCAGATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 901
Db 846 CTTTCCATAGCAGATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 905
Qy 902 CTTCAACCCAGCTTATAGAGCTATCCGCTTGAAGATTTTGTGCTACCATGATCATTA 961
Db 906 CTTCAACCCAGCTTATAGAGCTATCCGCTTGAAGATTTTGTGCTACCATGATCATTA 965
Qy 962 CTTGAGTGTGCTTATTAAGTGAAGCTGATGAAGATGATGAAGATGATGAAGATGAT 1021
Db 966 TTTACAGTTTGGTATTACAAGTTGAATGATCTGAAGATGATGAAGATGATGAAGAT 1025
Qy 1022 CTATCTGCTTCTTATCAGATGATCAAGGCTATGATGATGATGATGATGATGATGATGAT 1081
Db 1026 CTATCTGCTTCTTATCAGATGATCAAGGCTATGATGATGATGATGATGATGATGATGAT 1085
Qy 1082 CCACAGGAGACTTTGGGCTCTCCGATGAGCACTTCCAGCCCGAGCTTGAAGCACTTCA 1141
Db 1086 CCACAGGAGACTTTGGGCTCTCCGATGAGCACTTCCAGCCCGAGCTTGAAGCACTTCA 1145
```

```
Qy 1142 GTTCTTTAATGTGACCACTCCAGAGGAGCTGTATGACTTTGCTTAAGAAAATAATAAT 1201
Db 1146 GTTCTTTAATGTGACGAGCTCCAGAGGAACTGTACGACTTCGCTCAGGAACACCTAAT 1205
Qy 1202 GGATGATGATGAGGAGAGTGTGGAATATGTGATGATCTCTTGGAACTTGGAGGAGAC 1261
Db 1206 GGATGATGATGAGGAGAGTGTGGAATATGTGATGATCTCTTGGAGAGCGAAGAGTC 1265
Qy 1262 CAGCTAGCCACAGCAACCAAGAGACTTCT 1293
Db 1266 TGCCTAGTCCACAGGGCTAAGGAACCTTCT 1297
```

RESULT 13

AL5411167

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: segre@genoscope.cns.fr, Web: www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 2233.f For

more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CS0DE005DC06QPI&cluster=2233.f. Contact :

Feng Liang Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com/

Faraday Avenue Genoscope sequence ID: CS0DE005DC06QPI.

Location/Qualifiers

1. 1201

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DE005YF12"

/tissue_type="PLACENTA"

/clone_lib="Homo sapiens PLACENTA"

/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed

with a NotI-oligo(dT) primer. Five prime end enriched,

double-strand cDNA was digested with Not I and cloned into

the Not I and EcoRV sites of the pCMVSPORT 6 vector.

Library was not normalized."

BASE COUNT 258 a 335 c 313 g 237 t 58 others

ORIGIN

Query Match

Best Local Similarity

Matches

976; Conservative

10; Mismatches

4; Indels

4; Gaps

4;

Qy 57 GCGCGATGAGAGCG 116

Db 58 GCGCGATGAGAGCG 117

Qy 117 TCGCGCTGCG 176

Db 118 TCGCGCTGCG 176

Qy 177 GACGAGCTGATGCGCGCTGAGTTCGCGCTACCGGACCGCTGCGGCAAGTACAGCGGCGAG 236

Db 178 GACGAGCTGATGCGCGCTGAGTTCGCGCTACCGGACCGCTGCGGCAAGTACAGCGGCGAG 236

```
Db 177 GAGAGCTGATCGCGCTGAGTGGCTTACCGGCTACCGGCTGAGCAAGTACAGGGCGAG 236
QY 237 CACTGGCGGAGAGCTGGCTACCTGGAGATCAGCTCGGCTGCACCGCTTGTGCGC 296
Db 237 CACTGGCGGAGAGCTGGCTTACCTGGAGATCAGCTCGGCTGCACCGCTTGTGCGC 296
QY 297 GACAGGAGGCTTCTGCGCACCGCACTGACGCGCGCGCGCGCGCGCGCGCGCGC 356
Db 297 GACAGC-AGGCTTCTGCGCACCGCACTGACGCGCGCGCGCGCGCGCGCGCGCGC 355
QY 357 GCGCTCGCGAGTATCCCGAGTGGCGCTTCTGCGGGGCTGCTGCGCGCGCGCGCACTGC 416
Db 356 GCGCTCGCGAGTATCCCGAGTGGCGCTTCTGCGGGGCTGCTGCGCGCGCGCACTGC 415
QY 417 CTCAGCGCTGCAAGCAGGCGCTGCGAGCTTCTGCGCAGTCCAGCGCGCGCGCGAGTG 476
Db 416 CTCAGCGCTGCAAGCAGGCGCTGCGAGCTTCTGCGCAGTCCAGCGCGCGCGAGTG 475
QY 477 CTGGCGGACTTCAGGCGCGCGAGCGCTACAGTTCCTGCACTTCTGCTTACTTCAAGCA 536
Db 476 CTGGCGGACTTCAGGCGCGCGAGCGCTACAGTTCCTGCACTTCTGCTTACTTCAAGCA 535
QY 537 AATAATCTCCCAAGCCATGCGCGCTGCTCAACCTTCTACTGAAGCATCTGATGAC 596
Db 536 AATAATCTCCCAAGCCATGCGCGCTGCTCAACCTTCTACTGAAGCATCTGATGAT 595
QY 597 GAAATGATGAAGAGGAACATGATATTAAGAGCTGCGCTGCTGCGAGGACTACATT 656
Db 596 GAAATGATGAAGAGGAACATGATATTAAGAGCTGCGCTGCTGCGAGGACTACATT 655
QY 657 AAAGACCTGGAACCAAGTCATATGAAGCCTGTTTCATCCGAGCAGTGGCGGATACAA 716
Db 656 AAAGACCTGGAACCAAGTCATATGAAGCCTGTTTCATCCGAGCAGTGGCGGATACAA 715
QY 717 GGTGAGAACTGAGAACATCCATCAGACATGAGCTGGCGCTTCCCGACTTCTTCAA 776
Db 716 GGTGAGAACTGAGAACATCCATCAGACATGAGCTGGCGCTTCCCGACTTCTTCAA 775
QY 777 GCCTTTTACAGTGTCTCGAGCGCTCGAGGGTTCCAGGGAGATCAAGGACTTCAAGGAT 836
Db 776 GCCTTTTACAGTGTCTCGAGCGCTCGAGGGTTCCAGGGAGATCAAGGACTTCAAGGAT 835
QY 837 TTCTACCTTCCATAGACATATATGATGAGAGTCTCGAATGCAAAATACAGTGTGA 896
Db 836 TTCTACCTTCCATAGACATATATGATGAGAGTCTCGAATGCAAAATACAGTGTGA 895
QY 897 GAGAACTCACCCAGTATAGAGGCTATCCGGTTGAGAAATTTGTGGCTACCATGAT 956
Db 896 GAGAACTCACCCAGTATAGAGGCTATCCGGTTGAGAAATTTGTGGCTACCATGAT 955
QY 957 CATTACTTGCAGTTTGCCTATTATAAGTTGAACGACCTGAAGAATGACGCCCTGTGCA 1016
Db 956 CATTACTTGCAGTTTGCCTATTATAAGTTG-AMGACCTKGAAATGCAAG-CCCTGTGCA 1013
QY 1017 GTGAGTATCTGCTTCTTGTGATGAGTACAGG 1050
Db 1014 GTGAGTATCTGCTTCTTGTGATGAGTACAGG 1047

RESULT 14
AL572160/c 1074 bp mRNA linear EST 31-MAY-2003
LOCUS AL572160 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA.
DEFINITION clone CS0DI032YH24 3-PRIME, mRNA sequence.
ACCESSION AL572160
VERSION AL572160.2 GI:31293545
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1074)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
```

```
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT On Feb 16, 2001 this sequence version replaced gi:12930160.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 2233.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DI032DD12NP1&cluster=2233.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope, sequence ID : CS0DI032DD12NP1.
Location/Qualifiers
1. 1074
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI032YH24"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-cilgo(dt)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 325 a 204 c 283 g 237 t 25 others
ORIGIN
Query Match 23.9%; Score 924.6; DB 9; Length 1074;
Best Local Similarity 97.0%; Pred. No. 1.3e-12;
Matches 971; Conservative 10; Mismatches 16; Indels 4; Gaps 4;
QY 1307 AAACACAGATCTTTGTCCTTTCCACAGAGCCAGCGCTTGTATACCTCAGAGCCTTCT 1366
Db 999 AGAACACAGATCTTTTTCCTTTCCACAGAGCCAGCGCT-KTGATACCTCAGAGCCTTCT 941
QY 1367 CTTTACTCTCCAAAGTGAAGGAGGAGCCCGCTCTCTTAACCTGCATGTCATCAGGGTG 1426
Db 940 CTTTACTCTCCAAAGTGAAG-AGAAGCCCGCTCTCTTAACCTGCATGTCATCAGGGTG 892
QY 1427 AGCTGCTTCTCTATCTTCCACAGCTGCGACCTCATGTTTCACACCTATCTTTCTACCTT 1486
Db 881 AGCTGCTTCTCTATCTTCCACAGCTGCGACCTCATGTTTCACACCTATCTTTCTACCT-T 823
QY 1487 TTTTTCAGATGGAGTCTCGCTCTCTGCGCAGCTGGAGTGAATGGCAGCTTCTCAGC 1546
Db 822 TTTTTCAGATGGAGTCTCGCTCTCTGCGCAGCTGGAGTGAATGGCAGCTTCTCAGC 763
QY 1547 TCACCTGCAACCTCGCGCTCTTGGGTTCAAGCAATTCGTGCAATCAGCCTCCCGAGTACC 1606
Db 762 TCACCTGCAACCTCGCGCTCTTGGGTTCAAGCAATTCGTGCAATCAGCCTCCCGAGTACC 703
QY 1607 TGGGATTACAGGCAATGCGCACCGCGCTAAATTTTCTATTTTAGTAGAGAGCGGG 1666
Db 702 TGGGATTACAGGCAATGCGCACCGCGCTAAATTTTCTATTTTAGTAGAGAGCGGG 643
QY 1667 TTTTGCCATGTTGGCCAGGCTGCTCGAACTCTTGACTTCAGATGATCCATCGCCTTG 1726
Db 642 TTTTGCCATGTTGGCCAGGCTGCTCGAACTCTTGACTTCAGATGATCCATCGCCTTG 583
QY 1727 GCTTCCACAGTGTGGGATTACAGGCGTAGCCACCATGCGCGGCTCTTTCTCACCCTT 1786
Db 582 GCTTCCACAGTGTGGGATTACAGGCGTAGCCACCATGCGCGGCTCTTTCTCACCCTT 523
QY 1787 TACACCTGTCTTCTATCTTACCTCAGATCTGTTTTCACACCTTTCATCCCTGTCTTCTCATGT 1846
Db 522 TACACCTGTCTTCTATCTTACCTCAGATCTGTTTTCACACCTTTCATCCCTGTCTTCTCATGT 463
QY 1847 TCACACTGTCTTCCCATGTTTATAGCTGCTCTTCTTACCATTTGGTTTGAAGGCGAG 1906
Db 462 TCACACTGTCTTCCCATGTTTATAGCTGCTCTTCTTACCATTTGGTTTGAAGGCGAG 403
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2004, 08:05:34 ; Search time 218.243 Seconds
(without alignments)
7828.876 Million cell updates/sec

Title: US-09-729-674-1
Perfect score: 3871
Sequence: 1 tttctctctctctctctctt.....ataaaaaaaaaaaaaaaaaa 3871

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PTUS COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	466.4	12.0	618	3	US-09-328-111-215 Sequence 215, Appl
2	233.6	6.0	43950	4	US-09-735-934A-3 Sequence 3, Appli
3	233.6	6.0	43950	4	US-10-060-332-3 Sequence 3, Appli
4	229.4	5.9	55298	4	US-09-491-356C-1 Sequence 1, Appli
5	229	5.9	111282	4	US-09-754-250-3 Sequence 3, Appli
6	228.8	5.9	9365	4	US-09-608-285A-8 Sequence 8, Appli
7	228.8	5.9	9365	4	US-09-350-836B-8 Sequence 8, Appli
8	228.8	5.9	9365	4	US-09-370-265-8 Sequence 8, Appli
9	228.8	5.9	9365	4	US-09-557-800C-8 Sequence 8, Appli
10	228.8	5.9	14747	4	US-09-608-285A-42 Sequence 42, Appl
11	228.8	5.9	14747	4	US-09-557-800C-42 Sequence 42, Appl
12	228.8	5.9	15977	4	US-09-608-285A-59 Sequence 59, Appl
C 13	228.4	5.9	64467	4	US-09-803-671B-3 Sequence 3, Appli
14	227	5.9	11725	2	US-08-756-506-1 Sequence 1, Appli
C 15	226.6	5.9	55298	4	US-09-491-356C-1 Sequence 1, Appli
16	226.2	5.8	53332	4	US-09-801-861-3 Sequence 3, Appli
17	225.6	5.8	99500	4	US-09-798-096-10 Sequence 10, Appl
C 18	223.8	5.8	16063	4	US-09-801-053-3 Sequence 3, Appli
C 19	223.6	5.8	4773	3	US-08-884-324-9 Sequence 9, Appli
C 20	223.6	5.8	11464	3	US-08-884-324-13 Sequence 13, Appl
21	223.6	5.8	28994	3	US-08-884-324-14 Sequence 14, Appl
22	223.4	5.8	16389	4	US-09-741-154-3 Sequence 3, Appli
23	223	5.8	14581	4	US-08-520-373D-4 Sequence 4, Appli
24	223	5.8	22481	4	US-08-367-841A-43 Sequence 43, Appl
25	223	5.8	22481	5	PCT-US95-07201-43 Sequence 43, Appl
26	223	5.8	22484	4	US-09-875-223-2 Sequence 2, Appli
C 27	223	5.8	23187	4	US-09-499-522-1 Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-328-111-215/c
; Sequence 215, Application US/09328111
; Patent No. 6262333
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie.
; APPLICANT: Catino, Theodore J.
; APPLICANT: Derti, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; TITLE OF INVENTION: PRODUCTS
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/328,111
; EARLIER APPLICATION NUMBER: US 60/088,801
; EARLIER FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 215
; LENGTH: 618
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(618)
; OTHER INFORMATION: n = A,T,C or G
US-09-328-111-215

Query Match 12.0%; Score 466.4; DB 3; Length 618;
Best Local Similarity 94.8%; Pred. No. 2.ee-99;
Matches 506; Conservative 0; Mismatches 25; Indels 3; Gaps 3;
Qy 1076 CCAGTACCACAGGACACTTGGGGCTCTCGGATGAGCACTTCCAGCCAGACCTGAAGC 1135
Db 534 CNAAGTCCCNAGNCANTGGGGCTCTNGATGAGCANTTCCAGCCAGNCCTTAANC 475
Qy 1136 AGTTCAGTCTTTTAATGTGACCA-CACTCCAGAGAGGC-TGTATGACTTT-GCTAAGGA 1192
Db 474 AGTTCAGTCTTTAAATGTGNCACCACTCCAGANGAAGCTTGTATGACTTGGNTAAGGA 415
Qy 1193 AAATATATGATGATGATGAGGAGAGTCTGTGGATATGTGGATGACCTCTTGAAC 1252

Db 414 AATATATATGATGATGATGAGGAGAGAGTTGTGGAAATATGTGGATGACCTCTTGGAACT 355
QY 1253 GGAGGAGACAGTACAGCCACAGACACAAAGAGAGATTCCTCTTGGCGTTTCAGGAAACAC 1312
Db 354 GGAGGAGACAGTACAGCCACAGACACAAAGAGAGATTCCTCTTGGCGTTTCAGGAAACAC 295
QY 1313 AGATTCTTTGCTCTTTCCCAACAGCCAGGCTGTGTACCTCAGAGCTTCTCTTTAC 1372
Db 294 AGATTCTTTGCTCTTTCCCAACAGCCAGGCTGTGTACCTCAGAGCTTCTCTTTAC 235
QY 1373 TCTCCAAAGTGAAGGAAGAGCCCGCTCTCTTAAGTCAATGTCATCAGGGGTGAGCGTG 1432
Db 234 TCTCCAAAGTGAAGGAAGAGCCCGCTCTCTTAAGTCAATGTCATCAGGGGTGAGCGTG 175
QY 1433 CTTTCTCTATCTTCACACCTGCCACCTCAATGTTTCACACCTATCTTCTCACCTTTTTTTT 1492
Db 174 CTTTCTCTATCTTCACACCTGCCACCTCAATGTTTCACACCTATCTTCTCACCTTTTTTTT 115
QY 1493 GAGATGAGTCTCGCTCTCTTGGCCAGGCTGGAGTGCATGGAATGCAATGTCATCAGTCACTG 1552
Db 114 GAGATGAGTCTCGCTCTCTTGGCCAGGCTGGAGTGCATGGAATGCAATGTCATCAGTCACTG 55
QY 1553 CACCTCCCGCTCTTGGGTTCAAGCAATTCGTGTCATCAGCTCCCGAGTACC 1606
Db 54 CACCTCCCGCTCTTGGGTTCAAGCAATTCGTGTCATCAGCTCCCGAGTACC 1

RESULT 2

US-09-735-934A-3
; Sequence 3, Application US/09735934A
; Patent No. 6372468
; GENERAL INFORMATION:
; APPLICANT: LI, JIAYIN et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL000851
; CURRENT APPLICATION NUMBER: US/09/735,934A
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 43950
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-735-934A-3

Query Match 6.0%; Score 233.6; DB 4; Length 43950;
Best Local Similarity 86.8%; Pred. No. 4.1e-44;
Matches 257; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
QY 1478 TCTCACCTTTTTTTGAGATGAGTCTCGCTCTCTTGGCCAGGCTGGAGTGCATGGAAC 1537
Db 11765 TCTTTTTTTTTTTGAACTGAGTCTCGCTGTATTGCCAGGCTGGAGTGCAGTGGCGT 11824
QY 1538 GTTCTCAGCTCACTGCAACCTCCGCTCTTGGGTTCAAGCAATTCGTGTCATCAGCCCTC 1597
Db 11825 GATCTCAGCTCACTGCAACCTCTGCCTCTGGGTTCAAGTGTATCTCTGCTCAGCCTC 11884
QY 1598 CCGAGTACCTGGGATACAGGATGTCACCAACCGCCGGCTAAATTTGATTTTTAGTA 1657
Db 11885 CGGAGTACCTGGGATACAGGATGTCACCAACCGCCGGCTAAATTTGATTTTTAGTA 11944
QY 1658 GAGACGGGGTTTTGCCATGTTGCCAGGCTGTCTCGAACTCTTGACTTCAGATGATCCA 1717
Db 11945 GAGATGGGGTTTTGCCATGTTGCCAGGCTGTCTCGAACTCTTGACTTCAGATGATCCA 12004
QY 1718 TCTGCCTTGGCCTCCCAACAGTCTGGGATTCAGGGGTGAGCCACCATGCCCGGCC 1773
Db 12005 CCGGCTTGGCCTCCCAACAGTCTGGGATTCAGGGGTGAGCCACCATGCCCGGCC 12060

RESULT 3

US-10-060-332-3
; Sequence 3, Application US/10060332
; Patent No. 6528294
; GENERAL INFORMATION:
; APPLICANT: LI, JIAYIN et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL000851DIV
; CURRENT APPLICATION NUMBER: US/10/060,332
; CURRENT FILING DATE: 2002-02-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 43950
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-060-332-3

Query Match 6.0%; Score 233.6; DB 4; Length 43950;
Best Local Similarity 86.8%; Pred. No. 4.1e-44;
Matches 257; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
QY 1478 TCTCACCTTTTTTTGAGATGAGTCTCGCTCTCTTGGCCAGGCTGGAGTGCATGGAAC 1537
Db 11765 TCTTTTTTTTTTTGAACTGAGTCTCGCTGTATTGCCAGGCTGGAGTGCAGTGGCGT 11824
QY 1538 GTTCTCAGCTCACTGCAACCTCCGCTCTTGGGTTCAAGCAATTCGTGTCATCAGCCTC 1597
Db 11825 GATCTCAGCTCACTGCAACCTCTGCCTCTGGGTTCAAGTGTATCTCTGCTCAGCCTC 11884
QY 1598 CCGAGTACCTGGGATTCAGGATGTCACCAACCGCCGGCTAAATTTGATTTTTAGTA 1657
Db 11885 CGGAGTACCTGGGATTCAGGATGTCACCAACCGCCGGCTAAATTTGATTTTTAGTA 11944
QY 1658 GAGACGGGGTTTTGCCATGTTGCCAGGCTGTCTCGAACTCTTGACTTCAGATGATCCA 1717
Db 11945 GAGATGGGGTTTTGCCATGTTGCCAGGCTGTCTCGAACTCTTGACTTCAGATGATCCA 12004
QY 1718 TCTGCCTTGGCCTCCCAACAGTCTGGGATTCAGGGGTGAGCCACCATGCCCGGCC 1773
Db 12005 CCGGCTTGGCCTCCCAACAGTCTGGGATTCAGGGGTGAGCCACCATGCCCGGCC 12060

RESULT 4

US-09-491-356C-1
; Sequence 1, Application US/09491356C
; Patent No. 6566061
; GENERAL INFORMATION:
; APPLICANT: Philibert, Robert A.
; APPLICANT: Ginns, Edward I.
; APPLICANT: Delisi, Lynn
; TITLE OF INVENTION: IDENTIFICATION OF POLYMORPHISMS IN THE PCTG4 REGION OF XQ13
; FILE REFERENCE: 9465.6US11
; CURRENT APPLICATION NUMBER: US/09/491,356C
; CURRENT FILING DATE: 2000-01-26
; PRIOR APPLICATION NUMBER: PCT/US99/09365
; PRIOR FILING DATE: 1999-04-29
; PRIOR APPLICATION NUMBER: 60/083,465
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 55298
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (485)..(485)
; OTHER INFORMATION: n is not determined
; NAME/KEY: misc feature
; LOCATION: (836)..(838)
; OTHER INFORMATION: n is not determined

```

; NAME/KEY: misc feature
; LOCATION: (16728)..(16728)
; OTHER INFORMATION: n is not determined
; NAME/KEY: misc feature
; LOCATION: (22750)..(22750)
; OTHER INFORMATION: n is not determined
; NAME/KEY: misc feature
; LOCATION: (22756)..(22756)
; OTHER INFORMATION: n is not determined
; NAME/KEY: misc feature
; LOCATION: (28519)..(28519)
; OTHER INFORMATION: n is not determined
; NAME/KEY: misc feature
; LOCATION: (44804)..(44804)
; OTHER INFORMATION: n is not determined
; NAME/KEY: misc feature
; LOCATION: (45002)..(45002)
; OTHER INFORMATION: n is not determined
; NAME/KEY: misc feature
; LOCATION: (54049)..(54049)
; OTHER INFORMATION: n is not determined
; NAME/KEY: misc feature
; LOCATION: (54226)..(54226)
; OTHER INFORMATION: n is not determined
US-09-491-356C-1

Query Match      5.9%; Score 229.4; DB 4; Length 55298;
Best Local Similarity 80.7%; Pred. No. 4.4e-43;
Matches 280; Conservative 0; Mismatches 66; Indels 1; Gaps 1;

QY 1471 CTATCTTCTCACCTTTTTTTTGGAGATGGAGTCTCGTCTCTTCCCGCAGGCTGGAGTGCA 1530
Db 42875 CTTTCTTTTTTTTTTTTTTTTGGAGACAGATCTCGTCTCTCGCTCAGGCTGGAGTGCA 42934

QY 1531 ATGGACATCTCAGCTCACTGCAACCTCCCGCTCTTGGGTTCAAGCAATCTGCTGCAT 1590
Db 42935 ATGGCATGATCTCGCTCACTACACCTCCCGCTCCCGGTAAAGCAATCTCCTGCCT 42994

QY 1591 CAGCTCCCGAGTACCTGGATACAGCATGTGCCACCGCCGCGC-TAATTTTCTAT 1649
Db 42995 CAGCTCCCGAGTACCTGGATACAGCGGCTGCCACCGCCGCGCATTTTTGTAT 43054

QY 1650 TTTTAGTAGACAGCGGGTTTTGGCATGTGGCCAGGCTGGTCTCGAACTCTTGACTTCAG 1709
Db 43055 TTTTAGTAGACAGCGGGTTTACCATGTGGCCAGGCTGGTCTTGAACCTCGACCTCAG 43114

QY 1710 ATGATCCATCTGCTTGGCTCCACAGTCTGGGATACAGGCGTGAGCCACCATGCC 1769
Db 43115 GTGATCCACCTGCTCGGCTCCCAAGTGTGGGATACAGGCGATGAGCCATCACGCC 43174

QY 1770 GGCCTCTTCTCACCTTTTACACCTGTCTTCTATCTCTCACATCTGTT 1816
Db 43175 AGCCTTAAGGTGCTTTTCTAGATATCTTGGTGATTCATGAATGTT 43221

RESULT 5
US-09-754-250-3
; Sequence 3, Application US/09754250
; Patent No. 6376225
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHODIESTERASE
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: PHOSPHODIESTERASE PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001063
; CURRENT APPLICATION NUMBER: US/09/754,250
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 111282
; TYPE: DNA
; ORGANISM: Human

; NAME/KEY: misc feature
; LOCATION: (1)..(111282)
; OTHER INFORMATION: n = A,T,C or G
US-09-754-250-3

Query Match      5.9%; Score 229; DB 4; Length 111282;
Best Local Similarity 83.8%; Pred. No. 7.7e-43;
Matches 259; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 1480 TCACCTTTTTTTTGGAGATGGAGTCTCGTCTCTTCCCGCAGGCTGGAGTGCAATGCCAT 1539
Db 106289 TCTTGTTTTTTTTGGAGATGGAGTCTCACTCTGTTCCCGCAGGCTGGAGTGCAATGCCA 106348

QY 1540 TCTCAGCTCACTGCAACCTCCGCTCTTGGGTTCAAGCAATCTCTGCAATCAGCTCTCC 1599
Db 106349 TCTTGGCTCACTGCAACCTCCGCTCTTGGGTTCAAGCAATCTCTGCTCAGCTCTCC 106408

QY 1600 GAGTACCTGGGATTTACAGGCATGTGCCACCGCCGCTAAATTTTGTATTTTAGTAGA 1659
Db 106409 AAGTACCTGGGATTTACAGGCATGTGCCACCGCCGCTAAATTTTGTATTTTAGTAGA 106468

QY 1660 GACGGGGTTTTGCCATGTGGCCAGGCTGGTCTCGAACTCTTGAATTCAGATGCCATC 1719
Db 106469 GATGGGGCTTCACCAATTTTGGTCAGGCTGGTCTTGAATTCAGGTCATCCACA 106528

QY 1720 TGCCTTGGCTCCACAGTGTGGGATTTACAGGCGTGAGCCACCATGCCCGCTCTTTC 1779
Db 106529 CACCTCGGCATCTCGAGTGTGGGATTTACAGGCGTGAGCTACCGTACCTGGCCCTTGTG 106588

QY 1780 TCACCTTTA 1788
Db 106589 GAAATTTCA 106597

RESULT 6
US-09-608-285A-8
; Sequence 8, Application US/09608285A
; Patent No. 6335013
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 28110/36570
; CURRENT APPLICATION NUMBER: US/09/608,285A
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/583,231
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 8
; LENGTH: 9365
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3409)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
; NAME/KEY: misc feature
; LOCATION: (9214)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
; NAME/KEY: misc feature
; LOCATION: (9303)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
; NAME/KEY: misc feature
; LOCATION: (9311)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
US-09-608-285A-8

Query Match          5.9%; Score 228.8; DB 4; Length 9365;
Best Local Similarity 75.4%; Pred. No. 2.5e-43;
Matches 261; Conservative 34; Mismatches 47; Indels 4; Gaps 2;

QY 1462 TTTTCACACCTATCTTTCTCACCTTTTTCAGATGGAGTCTCGTC-TCTTGCCGAGG 1520
DB 7413 TATTATTATTATTATTATTATTATTATTATTGAGACGAGTCTGCTCTTGTTCCTYRGG 7472

QY 1521 CTGGAGTGAATGGACAGTCTCTCAGCTCACTGCAACCTCCGCTTGGGTTCAAGCAAT 1580
DB 7473 CTGGAGTGAATGGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7532

QY 1581 TCTGCTGATCAGCTCCGAGTACTCGGATTTACAGGATGTCACACCGCCGCGTA 1640
DB 7533 TCTGCTGATCAGCTCCGAGTACTCGGATTTACAGGATGTCACACCGCCGCGTA 7592

QY 1641 A---TTTTGTATTATTAGTAGAGAGCGGGTTTGGCATGTTGGCAGGCTGGTCTGAAC 1697
DB 7593 ATTTTGTATTATTATTAGTAGAGAGCGGGTTTACCAGTGTGGCAGGCTGCTCTYRAAC 7652

QY 1698 TCTGACTCAGATGATCATCTGCTTGGCTCCACAGTGGGATTTACAGGCGTGA 1757
DB 7653 TCTGACTCAGATGATCATCTGCTTGGCTCCACAGTGGGATTTAGAGGATTAAGGATGA 7712

QY 1758 GCCACATGCGCGGCTCTTTCTACCTTTACACCTGCTCTCTTAT 1803
DB 7713 GCCACATGCGCGGCTCTTTCTACCTTTACACCTGCTCTCTTAT 7758

RESULT 7
US-09-350-836B-8
; Sequence 8, Application US/09350836B
; Patent No. 6387645
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
; FILE REFERENCE: 28110/35761
; CURRENT APPLICATION NUMBER: US/09/350,836B
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 9365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
```

```
; LOCATION: (1)..(288)
; NAME/KEY: exon
; LOCATION: (1281)..(1580)
; NAME/KEY: exon
; LOCATION: (1820)..(1855)
; NAME/KEY: exon
; LOCATION: (2467)..(2555)
; NAME/KEY: exon
; LOCATION: (2863)..(2942)
; NAME/KEY: exon
; LOCATION: (3889)..(3950)
; NAME/KEY: exon
; LOCATION: (4894)..(4995)
; NAME/KEY: exon
; LOCATION: (5847)..(5987)
; NAME/KEY: exon
; LOCATION: (6966)..(7138)
; NAME/KEY: exon
; LOCATION: (8556)..(9365)
; NAME/KEY: misc feature
; LOCATION: (9409)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
; NAME/KEY: misc feature
; LOCATION: (9214)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
; NAME/KEY: misc feature
; LOCATION: (9303)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
; NAME/KEY: misc feature
; LOCATION: (9311)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
US-09-350-836B-8

Query Match          5.9%; Score 228.8; DB 4; Length 9365;
Best Local Similarity 75.4%; Pred. No. 2.5e-43;
Matches 261; Conservative 34; Mismatches 47; Indels 4; Gaps 2;

QY 1462 TGTTCACACCTATCTTTCTCACCTTTTTCAGATGGAGTCTCGTC-TCTTGCCGAGG 1520
DB 7413 TATTATTATTATTATTATTATTATTATTATTGAGACGAGTCTGCTCTTGTTCCTYRGG 7472

QY 1521 CTGGAGTGAATGGACAGTCTCTCAGCTCACTGCAACCTCCGCTTGGGTTCAAGCAAT 1580
DB 7473 CTGGAGTGAATGGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7532

QY 1581 TCTGCTGATCAGCTCCGAGTACTCGGATTTACAGGATGTCACACCGCCGCGTA 1640
DB 7533 TCTGCTGATCAGCTCCGAGTACTCGGATTTACAGGATGTCACACCGCCGCGTA 7592

QY 1641 A---TTTTGTATTATTAGTAGAGAGCGGGTTTGGCATGTTGGCAGGCTGGTCTGAAC 1697
DB 7593 ATTTTGTATTATTATTAGTAGAGAGCGGGTTTACCAGTGTGGCAGGCTGCTCTYRAAC 7652

QY 1698 TCTGACTCAGATGATCATCTGCTTGGCTCCACAGTGGGATTTACAGGCGTGA 1757
DB 7653 TCTGACTCAGATGATCATCTGCTTGGCTCCACAGTGGGATTTAGAGGATTAAGGATGA 7712

QY 1758 GCCACATGCGCGGCTCTTTCTACCTTTACACCTGCTCTCTTAT 1803
DB 7713 GCCACATGCGCGGCTCTTTCTACCTTTACACCTGCTCTCTTAT 7758

RESULT 8
US-09-370-265-8
; Sequence 8, Application US/09370265
; Patent No. 644771
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
; FILE REFERENCE: 28111/35908
; CURRENT APPLICATION NUMBER: US/09/370,265
```


SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 3
LENGTH: 64467
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(64467)
OTHER INFORMATION: n = A,T,C or G
US-09-803-671B-3

Query Match 5.9%; Score 228.4; DB 4; Length 64467;

Best Local Similarity 86.1%; Pred. No. 8.1e-43;

Matches 253; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY	1485	TTTTTTTGGAGTGGAGTCTCGCTCTCTTGCCAGGCTGGAGTGAATGGACGTTCTCA	1544
DB	24074	TTTTTTTGGAGCGAGGCTCACTCTGTCACCCAGTTGGAGTGAATGGTACAATCTCA	24015
QY	1545	GCTCATGCAACTCCGCTCTTGGGTTCAAGCAATCTGCTGCATCAGCCTCCCGAGTA	1604
DB	24014	GCTCATGCAACTCCGCTCCCGGTTCAACGATTTCTTGCTCAGCCTCCCGAGTA	23955
QY	1605	CTGGGATTACAGGATGTCACACAGCCCGGCTAAATTTGTATTTTGTAGTAGACGG	1664
DB	23954	GCTGGATTACAGTGGCGCACACACACCCGCTAAATTTGTATTTTGTAGTAGACGG	23895
QY	1665	GTTTTTGCATTTGGCCAGGCTGTCTGCAACTCTTGACTTCAGATCAATCTCGCT	1724
DB	23894	GTTTTTCAACATTTGTCAGGCTGTCTGAACTCTGACCTCAGTCAATCCGCCCACT	23835
QY	1725	TGGCTCCACAGTCTGGATTACAGCGTGAGCCACCATGCCCGGCTCTTT	1778
DB	23834	TGGCTCCCAAAGTCTGGATTACAGCGTGAGCCACATGTCCTCCGTCATTT	23781

RESULT 14

US-08-756-506-1
Sequence 1, Application US/08756506
Patent No. 5905185

GENERAL INFORMATION:

APPLICANT: Garner, Ian
APPLICANT: Cottingham, Ian R.
APPLICANT: Temperley, Simon M.
APPLICANT: Foster, Donald C.
APPLICANT: Sprecher, Cindy A.
APPLICANT: Prunkard, Donna E.
TITLE OF INVENTION: PROTEIN C PRODUCTION IN TRANSGENIC
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:

ADDRESSEE: ZymoGenetics, Inc.
STREET: 1201 Eastlake Avenue East
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98102

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/756,506
FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
NAME: Sawislak, Deborah A.
REGISTRATION NUMBER: 37,438
REFERENCE/DOCKET NUMBER: 95-28
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-442-6672
TELEFAX: 206-442-6678

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 11725 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: join(3520..3630, 5093..5117, 5210..5347, 5450
LOCATION: ..5584, 8253..8395, 9269..9386, 10516..11102)
US-08-756-506-1

Query Match 5.9%; Score 227; DB 2; Length 11725;

Best Local Similarity 85.0%; Pred. No. 7.2e-43;

Matches 266; Conservative 0; Mismatches 45; Indels 2; Gaps 1;

QY	1469	ACCTATCTTTCTCACCTTTTTTTTGAGATGGAGTCTCGCTCT--CTTGCCAGGCTGGAG	1526
DB <td>5361<th>ACTTTCTTTTCTTTTTTTTGAGATGGAGTTTCACTCTGTGTTGCCAGGCTGGAG</th><th>6420</th></td>	5361 <th>ACTTTCTTTTCTTTTTTTTGAGATGGAGTTTCACTCTGTGTTGCCAGGCTGGAG</th> <th>6420</th>	ACTTTCTTTTCTTTTTTTTGAGATGGAGTTTCACTCTGTGTTGCCAGGCTGGAG	6420
QY <th>1527</th> <th>TGCAATGGCAGCTTCTCAGCTCACTGCAACCTCCGCTCTTGGGTTCAAGCAATCTGCT</th> <th>1586</th>	1527	TGCAATGGCAGCTTCTCAGCTCACTGCAACCTCCGCTCTTGGGTTCAAGCAATCTGCT	1586
DB <th>6421</th> <th>TGCAATGACGTGATCTCAGCTCACCAACCTCCGCTCTCGATTCAGCGATTTCTCT</th> <th>6480</th>	6421	TGCAATGACGTGATCTCAGCTCACCAACCTCCGCTCTCGATTCAGCGATTTCTCT	6480
QY <th>1587</th> <th>GCATCAGCTCCGAGTACTGGATTACAGCATGTGCCACGCGCCGCTAATTTTG</th> <th>1646</th>	1587	GCATCAGCTCCGAGTACTGGATTACAGCATGTGCCACGCGCCGCTAATTTTG	1646
DB <th>6481</th> <th>GCCGAGCTCCGAGTACTGGATTACAGCATGTGCCACGCGCCGCTAATTTTG</th> <th>6540</th>	6481	GCCGAGCTCCGAGTACTGGATTACAGCATGTGCCACGCGCCGCTAATTTTG	6540
QY <th>1647</th> <th>TATTTTGTAGTAGAGCGGGTTTTTGCCATGTTGGCCAGGCTGCTCTCGAACTCTTGA</th> <th>1706</th>	1647	TATTTTGTAGTAGAGCGGGTTTTTGCCATGTTGGCCAGGCTGCTCTCGAACTCTTGA	1706
DB <th>6541</th> <th>TGTTTTGTAGTAGAAGGGTTTTTCTCGGTGTTGGTCAAGCTGCTCTTGAATCTCTG</th> <th>6600</th>	6541	TGTTTTGTAGTAGAAGGGTTTTTCTCGGTGTTGGTCAAGCTGCTCTTGAATCTCTG	6600
QY <th>1707</th> <th>CAGATGATCAATCTGCTTGCCCTCCACAGTCTGGGATTACAGCGTGGCCACCATG</th> <th>1766</th>	1707	CAGATGATCAATCTGCTTGCCCTCCACAGTCTGGGATTACAGCGTGGCCACCATG	1766
DB <th>6601</th> <th>CAGTATCACTGCTTGCCCTCTAAAGTCTGGGATTACAGCGTGGCCACCATG</th> <th>6660</th>	6601	CAGTATCACTGCTTGCCCTCTAAAGTCTGGGATTACAGCGTGGCCACCATG	6660
QY <th>1767</th> <th>CCCGGCTCTTTTCTTTC</th> <th>1779</th>	1767	CCCGGCTCTTTTCTTTC	1779
DB <th>6661</th> <th>CCGAGCTCTTTTCTTTC</th> <th>6673</th>	6661	CCGAGCTCTTTTCTTTC	6673

RESULT 15

US-09-491-356C-1/c
Sequence 1, Application US/09491356C
Patent No. 6566061
GENERAL INFORMATION:
APPLICANT: Philibert, Robert A.
APPLICANT: Ginns, Edward I.
APPLICANT: Delisi, Lynn
TITLE OF INVENTION: IDENTIFICATION OF POLYMORPHISMS IN THE PCTG4 REGION OF XQ13
FILE REFERENCE: 9465.6US11
CURRENT APPLICATION NUMBER: US/09/491,356C
CURRENT FILING DATE: 2000-01-26

PRIOR APPLICATION NUMBER: PCT/US99/09365

PRIOR FILING DATE: 1999-04-29

PRIOR APPLICATION NUMBER: 60/083,465

PRIOR FILING DATE: 1998-04-29

NUMBER OF SEQ ID NOS: 24

SOFTWARE: Patent in version 3.1

SEQ ID NO 1

LENGTH: 55298

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

LOCATION: (485)..(485)

OTHER INFORMATION: n is not determined

NAME/KEY: misc feature

LOCATION: (838)..(838)

OTHER INFORMATION: n is not determined

NAME/KEY: misc feature

LOCATION: (16728)..(16728)

OTHER INFORMATION: n is not determined
NAME/KEY: misc feature
LOCATION: (22750)..(22750)
OTHER INFORMATION: n is not determined
NAME/KEY: misc feature
LOCATION: (22756)..(22756)
OTHER INFORMATION: n is not determined
NAME/KEY: misc feature
LOCATION: (28519)..(28519)
OTHER INFORMATION: n is not determined
NAME/KEY: misc feature
LOCATION: (44804)..(44804)
OTHER INFORMATION: n is not determined
NAME/KEY: misc feature
LOCATION: (45002)..(45002)
OTHER INFORMATION: n is not determined
NAME/KEY: misc feature
LOCATION: (54049)..(54049)
OTHER INFORMATION: n is not determined
NAME/KEY: misc feature
LOCATION: (54226)..(54226)
OTHER INFORMATION: n is not determined

US-09-491-356C-1

Query Match 5.9%; Score 226.6; DB 4; Length 55298;
Best Local Similarity 83.8%; Pred. No. 2e-42;
Matches 269; Conservative 0; Mismatches 49; Indels 3; Gaps 1;

Qy	1471	CTATCTTCTCACCTTTTGTGAGATGGAGTCTCGTCTCTTGTCCCGAGCTGGAGTGCA	1530
Db	33384	CTTTTCTTTTCTTTTGTGAGATGGAGTCTCGTCTCTTGTCCCGAGCTGGAGTGCA	33325
Qy	1531	ATGGCAGCTTCTCAGCTCAGTCAACCTCCGCTCTTGGGTTCAAGCAATTCGTGCTCAT	1590
Db	33324	GTGGCAGATCTCGGCTCAGTCAAGCTCTGCTCTTGGGTTCAAGCAATTCGTGCTCAT	33265
Qy	1591	CAGCTTCCCGAGTACCTGGGATTCAGGCAATGTCACCCAGCTGGCTAAATTTGT---	1647
Db	33264	CAACCTCCCGAGTACCTGGGATTCAGGCAATGTCACCCAGCTGGCTAAATTTGT---	33205
Qy	1648	ATTTTGTAGAGACGGGTTTTCGCTATGTTGGCCAGCTGGTCTCGAATCTTGACTTC	1707
Db	33204	TTTTTGTAGAGACGGGTTTTCGCTATGTTGGCCAGCTGGTCTCGAATCTTGACTTC	33145
Qy	1708	AGATGATCCATCTGCTTGGCTCCCAAGTGTGGGATTCAGGCTGAGCCACCATGC	1767
Db	33144	AGGTGATCCATCTGCTTGGCTCCCAAGTGTGGGATTCAGGCTGAGCCACCATGC	33085
Qy	1768	CCGGCTCTTCTCACCTTTA	1788
Db	33084	CCGGCTCTCTCTGAGGTTTA	33064

Search completed: January 30, 2004, 18:32:22
Job time : 222.243 secs


```
QY 2221 CCGATCCCAATTGAGGACAATGCTTAGTTAATAGTCTCCGAGTTGGAAAAGGAAGGC 2280
Db 2221 CCGATCCCAATTGAGGACAATGCTTAGTTAATAGTCTCCGAGTTGGAAAAGGAAGGC 2280
QY 2281 CAGAGCTGTCTAGTATTCATTCATTCCTTTCAGTAAATATTTATTCAGTACCTACTGTGTGC 2340
Db 2281 CAGAGCTGTCTAGTATTCATTCATTCCTTTCAGTAAATATTTATTCAGTACCTACTGTGTGC 2340
QY 2341 TAGGCATTGACCTGGGAACCTAGAGATACCTCACAGAAATACAGGGAAGTTCCCTGTGCT 2400
Db 2341 TAGGCATTGACCTGGGAACCTAGAGATACCTCACAGAAATACAGGGAAGTTCCCTGTGCT 2400
QY 2401 CATGGAGCTTACATTTACAGGGAGAAAGAGATAGCCAAATACATAGGAATAAATATATAC 2460
Db 2401 CATGGAGCTTACATTTACAGGGAGAAAGAGATAGCCAAATACATAGGAATAAATATATAC 2460
QY 2461 AAGGTATCATGTAGTGATATTTCTGTGGAGAAAATAAGCAGGGGAGGAGTAAAGAA 2520
Db 2461 AAGGTATCATGTAGTGATATTTCTGTGGAGAAAATAAGCAGGGGAGGAGTAAAGAA 2520
QY 2521 TCCTGGAGATGAGCTGCAGTTTTAAATGGGGCCCTCACTGGGAATGTGACGTTGAGCAGA 2580
Db 2521 TCCTGGAGATGAGCTGCAGTTTTAAATGGGGCCCTCACTGGGAATGTGACGTTGAGCAGA 2580
QY 2581 GACGTTAGGGAACTGGATCCTGGACAGGCAATCCAGGCAGAGCAACAAGATGTGCACTG 2640
Db 2581 GACGTTAGGGAACTGGATCCTGGACAGGCAATCCAGGCAGAGCAACAAGATGTGCACTG 2640
QY 2641 CCCCAGAGTGAGAACTTGCTCTAGCTGTGTGAGAAAGAGCAGGGAGACCAAGCAGAGTCG 2700
Db 2641 CCCCAGAGTGAGAACTTGCTCTAGCTGTGTGAGAAAGAGCAGGGAGACCAAGCAGAGTCG 2700
QY 2701 TGGGAGGGGTAGAAATGGAAGAGAGGCGGCTGGGAGGACAGGTGTGGAGGCGCTTGG 2760
Db 2701 TGGGAGGGGTAGAAATGGAAGAGAGGCGGCTGGGAGGACAGGTGTGGAGGCGCTTGG 2760
QY 2761 CTTCTGCTTAAGTGTAGATGGGAACCACTGGAGGTTTGAAACAGAGGAGTGCCTGATTGAT 2820
Db 2761 CTTCTGCTTAAGTGTAGATGGGAACCACTGGAGGTTTGAAACAGAGGAGTGCCTGATTGAT 2820
QY 2821 TTATATTTGCAAGGGTCATCTAGTGCATATATGTGAAAACTTTAGTGGACAAGGCG 2880
Db 2821 TTATATTTGCAAGGGTCATCTAGTGCATATATGTGAAAACTTTAGTGGACAAGGCG 2880
QY 2881 AGAAGGAGGAGGAGACCTGTTAGGAGCTACTGCAAGGTTCCAGGCTTGGGCGCTGGC 2940
Db 2881 AGAAGGAGGAGGAGACCTGTTAGGAGCTACTGCAAGGTTCCAGGCTTGGGCGCTGGC 2940
QY 2941 CACAGCAACAGCAGTGGTCAATATCTAGATTTATTTTGAAGAGCCAATAGGATTTGC 3000
Db 2941 CACAGCAACAGCAGTGGTCAATATCTAGATTTATTTTGAAGAGCCAATAGGATTTGC 3000
QY 3001 TGAGAGTTTGAATGTGGAGTGTAGAGAGGAAAGATTAATGATGACATTAAGGTTTTTG 3060
Db 3001 TGAGAGTTTGAATGTGGAGTGTAGAGAGGAAAGATTAATGATGACATTAAGGTTTTTG 3060
QY 3061 GCCTGAATAGCAGGAAGATGGATACCGATTACTGAAATAGGGAAGGATGGGCTGGGT 3120
Db 3061 GCCTGAATAGCAGGAAGATGGATACCGATTACTGAAATAGGGAAGGATGGGCTGGGT 3120
QY 3121 AAGTAAGGAATTTGGTGAAGAGAGGCTGTCTGTGGTTGGAATGGGAGGTTCTGGCTGCA 3180
Db 3121 AAGTAAGGAATTTGGTGAAGAGAGGCTGTCTGTGGTTGGAATGGGAGGTTCTGGCTGCA 3180
QY 3181 AATCAAGCTGGAGATTTCTCTCAGGTCAGGTCAGTCTGAGCAGAGCTGAGACAGGATCTCAA 3240
Db 3181 AATCAAGCTGGAGATTTCTCTCAGGTCAGGTCAGTCTGAGCAGAGCTGAGACAGGATCTCAA 3240
QY 3241 TGCACCTTGGTTATTGTTGGGGTGTCTCTCAGAGGAACCTGTGAAAGGCTTTATCAGTC 3300
Db 3241 TGCACCTTGGTTATTGTTGGGGTGTCTCTCAGAGGAACCTGTGAAAGGCTTTATCAGTC 3300
```

```
QY 3301 ATTTATTCGCTGTGAGAAATTTCTCTGGAGTGTGGGTACATTTGAAGCAAGTGACTTCA 3360
Db 3301 ATTTATTCGCTGTGAGAAATTTCTCTGGAGTGTGGGTACATTTGAAGCAAGTGACTTCA 3360
QY 3361 GTTGAGGCAAGTCTCTGGAAGAGGCTGTAGGATCTGGCAGCTACCATCGTGGTAG 3420
Db 3361 GTTGAGGCAAGTCTCTGGAAGAGGCTGTAGGATCTGGCAGCTACCATCGTGGTAG 3420
QY 3421 TGTGTTGGGGGTGGGGGCTCTGGGCACCTGGCTGTGTGAAGGATCTGGCAGGGCACCACA 3480
Db 3421 TGTGTTGGGGGTGGGGGCTCTGGGCACCTGGCTGTGTGAAGGATCTGGCAGGGCACCACA 3480
QY 3481 GGGCCCCCTACTGAACCAATCAGCATGTCACTGGCATTTAAAGCCATGAGCTGGAGGGC 3540
Db 3481 GGGCCCCCTACTGAACCAATCAGCATGTCACTGGCATTTAAAGCCATGAGCTGGAGGGC 3540
QY 3541 CACTGAGATTGCTCTGAGTATTACTGAGAAACACAGAAAGAGCCATGGATGAGGCC 3600
Db 3541 CACTGAGATTGCTCTGAGTATTACTGAGAAACACAGAAAGAGCCATGGATGAGGCC 3600
QY 3601 TTGGGCTCTCTGGGAAATGGGAAATCAGCCAAAGGACTGAGAGGAGTTACCTTAAGGTC 3660
Db 3601 TTGGGCTCTCTGGGAAATGGGAAATCAGCCAAAGGACTGAGAGGAGTTACCTTAAGGTC 3660
QY 3661 AGAGAAAACCAAGAGAGTGTGTGTTCTGGAAGCTGAGCTTTCTTTATTTCAACCTCATTC 3720
Db 3661 AGAGAAAACCAAGAGAGTGTGTGTTCTGGAAGCTGAGCTTTCTTTATTTCAACCTCATTC 3720
QY 3721 CCTTCTCCAAATTAAGCCACTTTGTGTAGTTGGGCCCTCCAGGTTTGAAGGCAAGAGAGA 3780
Db 3721 CCTTCTCCAAATTAAGCCACTTTGTGTAGTTGGGCCCTCCAGGTTTGAAGGCAAGAGAGA 3780
QY 3781 AAGGCACAGCTTTGGGAAACAGAGACTTTCTGCAATAGCTGGGAAGGATTAAGAGA 3840
Db 3781 AAGGCACAGCTTTGGGAAACAGAGACTTTCTGCAATAGCTGGGAAGGATTAAGAGA 3840
QY 3841 TAGAGTGTATAAATAAAAAAATAAAAAA 3871
Db 3841 TAGAGTGTATAAATAAAAAAATAAAAAA 3871
```

RESULT 2

```
US-09-925-300-660
; Sequence 660, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: P101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 660
; LENGTH: 2549
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-300-660
```

Query Match 53.0%; Score 2051.2; DB 10; Length 2549;

Best Local Similarity 99.2%; Pred No. 0;
Matches 2077; Conservative 6; Mismatches 4; Indels 7; Gaps 2;

```
QY 1763 CATGCCCGGCTCTTTCTACCTTTACACCTGTCTTTATCTCTCATCTGTTTTCACA 1822
Db 206 CTGCGCGGCTCTTTCTACCTTTACACCTGTCTTTATCTCTCATCTGTTTTCACA 265
QY 1823 CCTTCATCCTCTCTCTCTCATGTTTTCACACTTGTCTTCCCATGTTTCATAGTGCCTTTC 1882
```


Db 266 CCTTCATCCCTGCTCTCTCATGTTACACATTTGCTTCCCATGTTTCATAGTGCCTTTC 325
Qy 1883 TTACCAATTTTGGTTGAAGGCGAGTCTTCTCTGGCTTGTCTTTTGGTTTTCCTCCAGAAAA 1942
Db 326 TTACCAATTTTGGTTGAAGGCGAGTCTTCTCTGGCTTGTCTTTTGGTTTTCCTCCAGAAAA 385
Qy 1943 TCAGTATTTATTTTAAATPAAGAAAAAATTCCTAGAGATGATAATTTGTGAAACCTCC 2002
Db 386 TCAGTATTTATTTTAAATPAAGAAAAAATTCCTAGAGATGAWAATTTGTGAAACCTCC 445
Qy 2003 TTTGGCTTATTTGCTTTTCCAGATTTTGTCTCTCTTTCCTCCATCCCGGAAAAAGATGTG 2062
Db 446 TTTGGCTTATTTGCTTTTCCAGATTTTGTCTCTCTTTCCTCCATCCCGGAAAAAGATGTG 505
Qy 2063 GAAGACATAGGCTAAATTTCTCCAGCTCCACAAATGCTTCTACTTGGTCTGACTTGTACC 2122
Db 506 GAAGACATAGGCTAAATTTCTCCAGCTCCACAAATGCTTCTACTTGGTCTGACTTGTACC 565
Qy 2123 AATTCCTAGACCCACTGAAAAAACAAGTTGAGTAGAGAGTGTAGAGTGCAGAAAAATGTGCT 2182
Db 566 AATTCCTAGACCCACTGAAAAAACAAGTTGAGTAGAGAGTGTAGAGTGCAGAAAAATGTGCT 625
Qy 2183 TTTGCCCACTTTGCTATCTCCAAATTTACAAAGTTGAGTGCAGTCCCATTTGAGGACAATG 2242
Db 626 TTTGCCCACTTTGCTATCTCCAAATTTACAAAGTTGAGTGCAGTCCCATTTGAGGACAATG 685
Qy 2243 CTTAGTTATAAGTCTCCGAGTTGAAAAAGGAAAAAGCCAGAGCTGTCTAGTTTCATTCA 2302
Db 686 CTTAGTTATAAGTCTCCGAGTTGAAAAAGGAAAAAGCCAGAGCTGTCTAGTTTCATTCA 745
Qy 2303 TTTCTTCAGTAATATTTTATTCAGTACCTTACTGTGTGTGTAGGATGACCTGG-----G 2356
Db 746 TTTCTTCAGTAATATTTTATTCAGTACCTTACTGTGTGTGTAGGATGACCTGGGAACTAG 805
Qy 2357 AACTAGAGTACTTTCACAGATAAACAAGGAGAAAGTTCCCTGTCTCATGGAGCTTACATTC 2416
Db 806 AACTAGAGTACTTTCACAGATAAACAAGGAGAAAGTTCCCTGTCTCATGGAGCTTACATTC 865
Qy 2417 TACAGGGAAGAGATAGCCAAATACATAGGAATAAATATATACAAAGTATCATGTAGTG 2476
Db 866 TACAGGGAAGAGATAGCCAAATACATAGGAATAAATATATACAAAGTATCATGTAGTG 925
Qy 2477 ATAAATTTCTGTGAGAAAAATAAAGCAGGGAGGAGTAAGAAATCCCTGGAGATGAGCT 2536
Db 926 ATAAATTTCTGTGAGAAAAATAAAGCAGGGAGGAGTAAGAAATCCCTGGAGATGAGCT 985
Qy 2537 GCAGTTTAAATGGGCTTCACTGGGAATGTGACGTTGAGCAGAGACGTTAGGGAAGTGG 2596
Db 986 GCAGTTTAAATGGGCTTCACTGGGAATGTGACGTTGAGCAGAGACGTTAGGGAAGTGG 1045
Qy 2597 ATCTGACAGAGCATTCAGGCGAGGAAACAAGATGTGCTGCTGCTCCCAAGTGAAGCT 2656
Db 1046 ATCTGACAGAGCATTCAGGCGAGGAAACAAGATGTGCTGCTGCTCCCAAGTGAAGCT 1105
Qy 2657 TGCTCTACGTTGTCAGGAAAGACAGGAGACCAAGCAGAGTCGTGGGAGGAGTGAAT 2716
Db 1106 TGCTCTACGTTGTCAGGAAAGACAGGAGACCAAGCAGAGTCGTGGGAGGAGTGAAT 1165
Qy 2717 GGAAGAGAGCGCGCTGGGAGAGACAGGTTGTGGAGGCGCTTGGCTTTCTGCTAAAGTGA 2776
Db 1166 GGAAGAGAGCGCGCTGGGAGAGACAGGTTGTGGAGGCGCTTGGCTTTCTGCTAAAGTGA 1225
Qy 2777 TGGGAACCACTGAGGAGGTTGAACAGAGGAGTGCCTTGTATTTATTTTGAAGGG 2836
Db 1226 TGGGAACCACTGAGGAGGTTGAACAGAGGAGTGCCTTGTATTTATTTTGAAGGG 1285
Qy 2837 TCATTTAGTGTCAATATTTGTGAAAAAATTTAGTGACAAGGCGACAGGAAGGGAAG 2896
Db 1286 TCATTTAGTGTCAATATTTGTGAAAAAATTTAGTGACAAGGCGACAGGAAGGGAAG 1345
Qy 2897 ACCTGTAGGAAGCTACTGCAAGTTCCAGGCTTGGGCTGGGCGACACACAGCAGTG 2956

Db 1346 ACCTGTTAGGAAGCTACTGCAAGGTTCCAGGCTTGGGCTGGGCCACAGCAACAGCAGTG 1405
Qy 2957 GTCAATATCTAGATTTATTTTGAAGAAGAGCCAAATAGGATTTGCTGAGAGTTGAATG 3016
Db 1406 GTCAATATCTAGATTTATTTTGAAGAAGAGCCAAATAGGATTTGCTGAGAGTTGAATG 1465
Qy 3017 GAGTGAAGAAAGAGAGGATTAATGATGACATTAAGAGTTTGGGCTTGAATAGCAGGAA 3076
Db 1466 GAGTGAAGAAAGAGAGGATTAATGATGACATTAAGAGTTTGGGCTTGAATAGCAGGAA 1525
Qy 3077 AGATGAGATTTACAGTTACTGAAATAGGAAGAGATGGCTGGTAAAGTAAAGNATTTGCT 3136
Db 1526 AGATGAGATTTACAGTTACTGAAATAGGAAGAGATGGCTGGTAAAGTAAAGNATTTGCT 1585
Qy 3137 GCAAGCAGGCTGCTGTGTGGTGGAAATGGAGGTTCTGGCTGCAAAATCAAAAGTGGAGA-T 3195
Db 1586 GCAAGCAGGCTGCTGTGTGGTGGAAATGGAGGTTCTGGCTGCAAAATCAAAAGTGGAGA-T 1645
Qy 3196 TCTCTCAGGTCAGGTCGAGCAGAGCTCGAGACAGGATCTGAATGCACTTGGTTTAT 3255
Db 1646 TCTCTCAGGTCAGGTCGAGCAGAGCTCGAGACAGGATCTGAATGCACTTGGTTTAT 1705
Qy 3256 GTTGGGGTCTCTCAGAGAGACCTGTGAAAGCCCTTTATCAGTCACTTTATTTGGCTGTGA 3315
Db 1706 GTTGGGGTCTCTCAGAGAGAACCTGTGAAAGCCCTTTATCAGTCACTTTATTTGGCTGTGA 1765
Qy 3316 GAAATTTCTTGGAGTGTGGTACATTTGAAGCAAGTGAATTCAGTCACTTTATTTGGCTGTGA 3375
Db 1766 GAAATTTCTTGGAGTGTGGTACATTTGAAGCAAGTGAATTCAGTCACTTTATTTGGCTGTGA 1825
Qy 3376 CTGGAAGAGGCTGTAGGCTCTGCAGCTACCATCGTGGTGTGTGTGGGGTGG 3435
Db 1826 CTGGAAGAGGCTGTAGGCTCTGCAGCTACCATCGTGGTGTGTGTGGGGTGG 1885
Qy 3436 GGTCTCTGGGCACTGGTGTGTGAAGGATCTTGGCAGGCAACCAAGCCCTTCTCTGGA 3495
Db 1886 GGTCTCTGGGCACTGGTGTGTGAAGGATCTTGGCAGGCAACCAAGCCCTTCTCTGGA 1945
Qy 3496 CCATCAGCATGTCACTGGGCACTTTAAAGCCATGAGCTGGAGGGCCCTGAGNATTTCTC 3555
Db 1946 CCATCAGCATGTCACTGGGCACTTTAAAGCCATGAGCTGGAGGGCCCTGAGNATTTCTC 2005
Qy 3556 TGAGTATTACTGAGAACCAAGAGCCATGATGAGCCCTTGGGCTCTCTGGA 3615
Db 2006 TGAGTATTACTGAGAACCAAGAGCCATGATGAGCCCTTGGGCTCTCTGGA 2065
Qy 3616 AATGGGAATCAGCCAAAGACTGAGAGAGTTACCTTAAGGTGAGAAACCAAGAG 3675
Db 2066 AATGGGAATCAGCCAAAGACTGAGAGAGTTACCTTAAGGTGAGAAACCAAGAG 2125
Qy 3676 AGTGTGTGTCTGGAGCTGAGCTTTCTTATTCAACCTCATTCCTTCTCCAAATAAG 3735
Db 2126 AGTGTGTGTCTGGAGCTGAGCTTTCTTATTCAACCTCATTCCTTCTCCAAATAAG 2185
Qy 3736 CCATTTGTGTGTGGGCGCTCCAGGTTGAAGCAAGAGGAGAAAGCAGAGCTTTG 3795
Db 2186 CCATTTGTGTGTGGGCGCTCCAGGTTGAAGCAAGAGGAGAAAGCAGAGCTTTG 2245
Qy 3796 GGAACAGAGCTTTCTGCAATAGCTGGAGGATTAAGGATAGAGTGT 3849
Db 2246 GGAACAGAGCTTTCTGCAATAGCTGGAGGATTAAGGATAGAGTGT 2299

RESULT 3

US-10-108-260A-2058
; Sequence 2058, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108, 260A
; CURRENT FILING DATE: 2002-03-27

NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1.
; SEQ ID NO 2058
; LENGTH: 1810
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-2058

Query Match 42.8%; Score 1656; DB 12; Length 1810;

Best Local Similarity 93.1%; Pred. No. 0;
Matches 1806; Conservative 2; Mismatches 2; Indels 130; Gaps 2;

Qy	39	CCTTCCTTCTCTTCCCGGCGCGATGAGACCGGGGCGCGGGCGCGCGCTGCTA	98
Db	1	CCTTCCTTCTCTTCCCGGCGCGATGAGACCGGGGCGCGGGCGCGCGCTGCTA	60
Qy	99	GGCGTGTGTGGTGGCTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG	158
Db	61	GGCGTGTGTGGTGGCTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG	120
Qy	159	TTCCGACGCTTCCACCGGACGAGCTGATGCGGCTCGAGTCGGCTACCGSCACGCGCTG	218
Db	121	TTCCGACGCTTCCACCGGACGAGCTGATGCGGCTCGAGTCGGCTACCGSCACGCGCTG	180
Qy	219	GACAGTACAGCGGACGACTGGGCCGAGAGCTKGGCTACCTGGAGTACAGCTGGG	278
Db	181	GACAGTACAGCGGACGACTGGGCCGAGAGCTGGGCTACCTGGAGTACAGCTGGG	240
Qy	279	CTGCACCGCTTGTCTCGCGGACAGCAGAGGCTTCTGCCACCGCACTGCGCGCGCGCG	338
Db	241	CTGCACCGCTTGTCTCGCGGACAGCAGAGGCTTCTGCCACCGCACTGCGCGCGCGCG	300
Qy	339	CAGCCGAGCGCGCGCGCGCTCGCAGCTATCCCGAGCTTCGCGCTTTCGGGGCGCTG	398
Db	301	CAGCCGAGCGCGCGCGCGCTCGCAGCTATCCCGAGCTTCGCGCTTTCGGGGCGCTG	360
Qy	399	GTGCGCGCGCGCACTGTCTCAAGCGCTGCAAGCAGAGGCGCTGCCAGCGCTTCCGCGAGTCC	458
Db	361	GTGCGCGCGCGCACTGTCTCAAGCGCTGCAAGCAGAGGCGCTGCCAGCGCTTCCGCGAGTCC	420
Qy	459	CAGCCGAGCGGAGTGTGCGGACCTTCAGCGCTTCAGCGCGCGCGCGCGCTCAAGTTCTCGAG	518
Db	421	CAGCCGAGCGGAGTGTGCGGACCTTCAGCGCTTCAGCGCGCGCGCGCGCTCAAGTTCTCGAG	480
Qy	519	TTGCTTACTCAAGGCAATTAATCTCCCAAGCAATCGCGCTTCACACCTTTCTA	578
Db	481	TTGCTTACTCAAGGCAATTAATCTCCCAAGCAATCGCGCTTCACACCTTTCTA	540
Qy	579	CTGAAGCATCTGTGACGAAATGATGAAGAGGAACATGGCATATTATAGAGCGCTGCCT	638
Db	541	CTGAAGCATCTGTGACGAAATGATGAAGAGGAACATGGCATATTATAGAGCGCTGCCT	600
Qy	639	GGTCCGAGGACTACATTAAAGACCTGGAACCAAGTATATGAAGCGTTTCATCCGA	698
Db	601	GGTCCGAGGACTAC-ATTAAAGCCTGGAACCAAGTATATGAAGCGTTTCATCCGA	659
Qy	699	GCAGTGCGGGATACAAACGGTGAACCTGGAGAACATCCATCACAGACATGGAGCTGGCC	758
Db	660	GCAGTGCGGGATACAAACGGTGAACCTGGAGAACATCCATCACAGACATGGAGCTGGCC	719
Qy	759	CTTCCGAGCTTCTCAAGGCTTTTACAGTGTCTCGAGCTGCTCGAGCGTTCAGGGAG	818
Db	720	CTTCCGAGCTTCTTCAAGGCTTTTACAGTGTCTCGAGCTGCTCGAGCGTTCAGGGAG	779
Qy	819	ATCAAGGACTTCAAGGATTTCTACCTTCCATAGCAGATCATTTATGATAGAGTTCTCGAA	878
Db	780	ATCAAGGACTTCAAGGATTTCTACCTTCCATAGCAG- - - - -	816
Qy	879	TGCAAAATACAGTGTGAAGAGAACTCACCCAGTTATAGAGGCTATCCGGTTGAGAAA	938
Db	817	-----	816
Qy	939	TTTGTGGCTACCATGATCATTTACTTGTGAGTTTGTGCTATTATATAGTTGAAGCGCTGAAG	998

RESULT 4

Db	817	-----TGACGACCTGAAG	830
Qy	999	AATGAGCGCCCTGTGAGTCACTATCTCTTTGATCAGAAATGACAGCTCATGCG	1058
Db	831	AATGAGCGCCCTGTGAGTCACTATCTCTTTGATCAGAAATGACAGCTCATGCG	890
Qy	1059	CAGAACCTGTGTATTATACAGGACACACAGGACACTTGGGGCTCTCGGATGAGCACTTC	1118
Db	891	CAGAACCTGTGTATTATACAGGACACACAGGACACGTTGGGGCTCTCTCAGATGAGCACTTC	950
Qy	1119	CAGCCAGACCTGAGCAGCTTCACTTTTATATGACACACACTCCAGAAAGAGCTGTAT	1178
Db	951	CAGCCAGACCTGAGCAGCTTCACTTTTATATGACACACACTCCAGAAAGAGCTGTAT	1010
Qy	1179	GACTTTGTCTAAGGAAATATATGATGATGATGAGGAGAAAGTTGTGGAAATATGTGAT	1238
Db	1011	GACTTTGTCTAAGGAAATATATGATGATGATGAGGAGAAAGTTGTGGAAATATGTGAT	1070
Qy	1239	GACCTCTTGGAACTGGAGGAGACCACTAGCCACAGCAACCAAGAGACTTCTCTCTGG	1298
Db	1071	GACCTCTTGGAACTGGAGGAGACCACTAGCCACAGCAACCAAGAGACTTCTCTCTGG	1130
Qy	1299	CGTTCAAGGAAACACAGATTTCTTTGCTTTTCCCAACAGCCAGAGCTGTGATACCTCAG	1358
Db	1131	CGTTCAAGGAAACACAGATTTCTTTGCTTTTCCCAACAGCCAGAGCTGTGATACCTCAG	1190
Qy	1359	AGCCTTCTCTTTACTCTCCAAAGTGAAGGAGAGCCCGCTCTCTCTAACTCATGTCTAT	1418
Db	1191	AGCCTTCTCTTTACTCTCCAAAGTGAAGGAGAGCCCGCTCTCTCTAACTCATGTCTAT	1250
Qy	1419	CAGGGGTGAGCTGCTCTTCTCTATCTTACACCTGCCACCTCATGTTACACCTATCTTT	1478
Db	1251	CAGGGGTGAGCTGCTCTTCTCTATCTTACACCTGCCACCTCATGTTACACCTATCTTT	1310
Qy	1479	CTCACCTTTTTTTTGAGATGAGTCTCGCTCTCTTTGCCAGGCTGGAGTGAATGGCAGC	1538
Db	1311	CTCACCTTTTTTTTGAGATGAGTCTCGCTCTCTTTGCCAGGCTGGAGTGAATGGCAGC	1370
Qy	1539	TTCTGAGTCACTGCAACCTCGGCTCTTGGGTTCAAGCAATCTGCTGATCAGCTCC	1598
Db	1371	TTCTGAGTCACTGCAACCTCGGCTCTTGGGTTCAAGCAATCTGCTGATCAGCTCC	1430
Qy	1599	CGAGTACCTGGGATTACAGGCAATGTGCCACACCGCCGCTAAATTTGTATTTTAGTAG	1658
Db	1431	CGAGTACCTGGGATTACAGGCAATGTGCCACACCGCCGCTAAATTTGTATTTAGTAG	1490
Qy	1659	AGACGGGGTTTTGCCATGTTGCCAGGCTGTCTCGAACTCTTGACTTCAGATGATCCAT	1718
Db	1491	AGACGGGGTTTTGCCATGTTGCCAGGCTGTCTCGAACTCTTGACTTCAGATGATCCAT	1550
Qy	1719	CTGCTTGGCTCCCGACAGTGTGGGATTACAGGCGTGAGCCACCATGCCCGGCTCTTT	1778
Db	1551	CTGCTTGGCTCCCGACAGTGTGGGATTACAGGCGTGAGCCACCATGCCCGGCTCTTT	1610
Qy	1779	CTCACCTTTTACACCTGTCTTCTTATCCCTCACATCTGTTTTTACACCTTCATCCCTGTCT	1838
Db	1611	CTCACCTTTTACACCTGTCTTCTTATCCCTCACATCTGTTTTTACACCTTCATCCCTGTCT	1670
Qy	1839	CCTCATGTTTACACTGTCTTCCCATGTTTCACTGCTCTTCTTACCATTTTGGTTG	1898
Db	1671	CCTCATGTTTACACTGTCTTCCCATGTTTCACTGCTCTTCTTACCATTTTGGTTG	1730
Qy	1899	AAGGGCAGTCTTCTGCTGCTTTTTTTTTTCCAGAGAAATCAGTATTATTTTTTTA	1958
Db	1731	AAGGGCAGTCTTCTGCTGCTTTTTTTTTTCCAGAGAAATCAGTATTATTTTTTTA	1790
Qy	1959	AATAAGAAAAATTCCTTAG	1978
Db	1791	AATAAGAAAAATTCCTTAG	1810

Qy	652	ACATTAAGACCTGGAAACCAAGTCATATGAAGCCTGTGTTATCCGAGCAGTGCGGCAT	711
Db	1051	CCCTCAGCGACCTAGAGGCCAGCCCTACAGGCCGTGTCTCTCGGCGCTGTGAAGCTCT	1110
Qy	712	ACACGGTGAGAACTGGAGAACATCCATCACAGCATGGAGCTGGCCCTCCCGACTTCT	771
Db	1111	ACACACGGGGGATTTCCGCAGCAGCAGGAGACATGGAGCGGCGCTTCTCAGAGTACC	1170
Qy	772	TCAAAGCCTTTTACGAGTGTCTCGCAGCCTCGCAGGGTTCCAGGGAGATCAAGGACTTCA	831
Db	1171	TGGCAGTCTTTCGCGGTGCTGGCGGCTGTGAAGGGGCCATGACGAGGTGGACTTCA	1230
Qy	832	AGGATTTCTACCTTTCCATAGCAGATCAATTATGTAGAAGTTCGGAAATGCAAAATACAGT	891
Db	1231	AGGACTTCTACCGGCCATAGCAGATCTCTTTGCAGAGTCCCTGCAGTCAAGGTGGACT	1290
Qy	892	GTGAGAGAACCTCACCCAGTTATAGGAGGCATCCGGTTGAGAAATTTGTGGCTACCA	951
Db	1291	GTGAGGCCAATTTGACCCCCAATGTGGGTGGCTACTTCTGTGGACAAGTTCGTGGCCACCA	1350
Qy	952	TGTATCATTTACTTGCAGTTTGCCTTATTTAAGTTGAACGACCTGAAGAAATGCAGCCCCCT	1011
Db	1351	TGTACCACTACCTGCAGTTTGCCTACTATTAAGTTGATGATGTGGCCAGGCTGCCGCA	1410
Qy	1012	GTGCAGTCACTATCTCCTCTTTTGATCAGAAATGACAAGGTCAATGCAGCAGAACTGGTGT	1071
Db	1411	GCGCCGCCAGCTACATGCTCTTCGACCCCAAGGACAGCGTCATGCAGCAGAACTGGTGT	1470
Qy	1072	ATTACCAGTACCACAGGACACTTGGGGGCTCTCGGATGAGCACTTCAGCCCCAGACCTG	1131
Db	1471	ATTACCGGTTCCACCGGGCTGCTGGGGCTGGGAAGGAGGAGTCTCCAGCCCGGGAGG	1530
Qy	1132	AAGCAGTTCAGTTCTTTAATGTGACCACACTCCAGAAGGAGCTGTATGACTTTGCTAAGG	1191
Db	1531	AGGCCATGCTTACCACAAACGAGNCCGCCGAGCTGCGGGAGCTGCTGGAGTTTACCCACA	1590
Qy	1192	AAATAATAATGGATGATGATGAGGGAGAAGTTTGGGA	1228
Db	1591	TGTACCTGCGAGTCAGATGATGAGATGAGCTGGAGGA	1627

RESULT 6

US-09-879-536-215/c
; Sequence 215, Application US/09879536
; Patent No. US2002014298A1
; GENERAL INFORMATION:

```

? GENERAL INFORMATION:
?
? APPLICANT: Endege, Wilson O.
? APPLICANT: Steinmann, Kathleen E.
? APPLICANT: Astle, Jon H.
? APPLICANT: Burgess, Christopher C.
? APPLICANT: Bushnell, Steven E.
? APPLICANT: Carroll III, Eddie
? APPLICANT: Catino, Theodore J.
? APPLICANT: Derti, Adnan
? APPLICANT: Ford, Donna M.
? APPLICANT: Lewis, Marcia E.
? APPLICANT: Monahan, John E.
? APPLICANT: Schlegel, Robert
?
? TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
?
? FILE REFERENCE: CCD-257 (US)
?
? CURRENT APPLICATION NUMBER: US/09/879,536
?
? CURRENT FILING DATE: 2001-09-21
?
? PRIOR APPLICATION NUMBER: US 60/088,801
?
? PRIOR FILING DATE: 1998-06-10
?
? NUMBER OF SEQ ID NOS: 850
?
? SOFTWARE: FastSeq for Windows Version 3.0

```

; LOCATION: (1)...(618)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-879-536-215

Query Match	12.0%; Score 466.4; DB 10; Length 618;
Best Local Similarity	94.8%; Pred. No. 6.5e-116;
Matches 506; Conservative 0; Mismatches 25; Indels 3; Gaps 3;	
QY	1076 CAGTACCACAGGACACTGGGGCCTCTCGGATGACACTCCAGCCAGACCTGAAC 1135
DB	534 CNAAGTCCNNAAGNCANTGGGCGCTNTGAATGACANNTCCAGCCAGNCTTAAAC 475
QY	1136 AGTTCAGTCTCTTAATGTGACCA-CACTCAGAAGGAGC-TGTATGACTTT-GCTAAGGA 1192
DB	474 AGTTCAGTCTCTTAATGTGNCACCACTCCAGANGAAGCTTGTATGACTTGGNTAAGGA 415
QY	1193 AAATATAATGGATGATGATGAGGGAAGTGTGGAAATATGTGGATGACCTCTTGGAACT 1252
DB	414 AAATATAATGGATGATGATGAGGGAAGTGTGGAAATATGTGGATGACCTCTTGGAACT 355
QY	1253 GGAGGAGACAGCTAGCCACACAGCCAGGCTGTTGATACCTCAGAGCCTTCTCTTTAC 1312
DB	354 GGAGGAGACAGCTAGCCACACAGCCAGGCTGTTGATACCTCAGAGCCTTCTCTTTAC 295
QY	1313 AGATTCCTTTGTCCTTTTCCCAACAGCCAGGCTGTTGATACCTCAGAGCCTTCTCTTTAC 1372
DB	294 AGATTCCTTTGTCCTTTTCCCAACAGCCAGGCTGTTGATACCTCAGAGCCTTCTCTTTAC 235
QY	1373 TCTCCAAAGTGAAGGGAAGCCCCCTCTCTCTAACTGCAATGATCAGGGGTGAGCCTG 1432
DB	234 TCTCCAAAGTGAAGGGAAGCCCCCTCTCTCTAACTGCAATGATCAGGGGTGAGCCTG 175
QY	1433 CTTTTCCTATCTTCACACCTGCCACCTCATGTTACACCTATCTTCTCACCCTTTTTTT 1492
DB	174 CTTTTCCTATCTTTCACACCTGCCACCTCATGTTACACCTATCTTCTCACCCTTTTTTT 115
QY	1493 GAGATGGAGTCTCGCTCTCTTGGCCAGGCTGGAGTGCATAGGCAACGTTCTCAGCTCACTG 1552
DB	114 GAGATGGAGTCTCGCTCTCTTGGCCAGGCTGGAGTGCATAGGCAACGTTCTCAGCTCACTG 55
QY	1553 CAACCTCCGCTCTTGGGTTCAAGCAATTCGTGTCATCAGCTCCCGATACC 1606
DB	54 CAACCTCCGCTCTTGGGTTCAAGCAATTCGTGTCATCAGCTCCCGATACC 1

RESULT 7

```

RESULTS:
US-09-962-832-113
; Sequence 113, Application US/09962832
; Patent No. US20020110821A1
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Cancer Gene Determina
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-74
; CURRENT APPLICATION NUMBER: US/09/962,832
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,077
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,280
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 259
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 113
; LENGTH: 2347
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-832-113

```

Query Match 12.0%; Score 465.8; DB 10; Length 2347;
Best Local Similarity 64.2%; Pred. No. 2.2e-115;
Matches 752; Conservative 2; Mismatches 399; Indels 18; Gaps 3;

Db	601	CGGACCTAGAGCCCGAGCCCTACGAGGCGGTCTCCCTCCGGCTGTGAAGCTCTACAACA	660
Qy	718	GTGAGAACTGGAGAACTCCATCACAGACATGGAGCTGGCCCTTCCCGACTCTTCAAG	777
Db	661	CGGGGATTTCCGACAGCAGCGAGGACATGGAGCGGCTTGTTCAGAGTACCTGGCAG	720
Qy	778	CCCTTTACAGTGTCTCCAGCCTCGAGGGTTCCAGGAGATCAAGACTTCAAGATT	837
Db	721	TCCTTGGCCCGTGCCTGCCCGCTGTGAAGGGGCCATGACGAGGTGGACTTCAAGACT	780
Qy	838	TCTACTTTCCTATAGCAGATCAATTATGAPAGAAGTCTCGAATGCAAAATACAGGTGCAAG	897
Db	781	TCTACCCGGCCATAGCAGATCTCTTTGCAGAGTCCCTGCACTGCAAGTGGACTGTGAGG	840
Qy	898	AGAACCTCACCCAGTTATAGAGGCTATCCGGTTGAGAAATTTGTGGCTACCATGTATC	957
Db	841	CCAAATTCACCCCAATGTGGTGGCTACTTCGTGGACAAGTTCGTGGCCACCATGTACC	900
Qy	958	ATTACTTCAGTTTGCCTATTATTAAGTTGAACGACCTGAAGAATGACAGCCCCCTGTGCAG	1017
Db	901	ACTACTCGAGTTTGCCCTACTATTAAGTTGAATGATGTGCGCCAGGCTGCCGAGCGCGG	960
Qy	1018	TCAGTATCTGCTCTTTGATCAGAATGACAAGTGCATGCAGCAAACTGGTGTATTACC	1077
Db	961	CCAGCTACATGCTCTTCGACCCCAAGGACACGCTCATGCAGCAAACTGGTGTATTACC	1020
Qy	1078	AGTACCACAGGACACTTGGGGCTCTCGATGAGACATTCACGCCAGACCTGGAAGCAG	1137
Db	1021	GGTTCCACCGGGCTCGCTGGGGCTGGAAGAGGAGGACTTCACGCCCGGAGGAGGCCA	1080
Qy	1138	TTCAAGTTCTTTAATGTGACCACATCCAGAAAGGAGCTGTATGCTTTGCTAAGGAAAATA	1197
Db	1081	TGCTCTACCACAAACAGACCGCCGAGCTGCGGGAGCTGTGGAGTTTCAACCCATGTACC	1140
Qy	1198	TAATGGATGATGATGAGGGAGAACTTTGGA	1228
Db	1141	TGCACTCAGATGATGAGATGAGCTGGAGGA	1171

```

RESULT 9
; US-09-918-995-22977
; Sequence 22977, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22977
; LENGTH: 494
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-918-995-22977

```

QY	3379	GAAGAAGGCTGTAGGCATCTGGCAGCTACCATGCGTGGTAGTGTGTGGGGGTGGGGGT	3439
Db	151	GAAGAAGGCTGTAGGCATCTGGCAGCTACCATGCGTGGTAGTGTGTGGGGGTGGGGGT	210
QY	3439	CTTGGGCACCTGGCTGTGTGTGAAGGATCTGGCAGGGGACCAACAGCGGCCCTTACTGAACCA	3498
Db	211	CTTGGGCACCTGGCTGTGTGTGAAGGATCTGGCAGGGGACCAACAGCGGCCCTTACTGAACCA	270
QY	3499	TCAGCATGTCAGTGGCATTTTAAAGCCATGCGAGCTGGAGGGGCCACCTAGAGATTGTCCTGA	3558
Db	271	TCAGGATGTCAGTGGCATTTTAAAGCCATGCGAGCTGGAGGGGCCACCTAGAGATTGTCCTGA	330
QY	3559	GTATTACTGAGAAGCAACAGAAAAGACCCATGATGGAGCCCTTGGGCTCTCTGGGAAAT	3618
Db	331	GTATTACTGAGAAGCATCAGATAGAGCCATGATGGAGCCCTTGGGCTCTCTGGGAAAC	390
QY	3619	GGGAATCAGCCAAAGGACTCGAAGGAGTTACCTTAAAGTCAGAGAAACCAAGAGCT	3678
Db	391	GGGAATCAGCCATAGGACTCGAAGGAGCTACCTTACAGTCAGAGAAACCAATAGAGT	450
QY	3679	GTGTGTGTCGGAAGCTGAGCTTCTTTTATTCAACCTCATTTCCC	3722
Db	451	GTGGCGCCTGGAACTGAGCTTCTTTATATAAACCTCATTTCTC	494

RESULT 10
 US-09-880-107-1453/c
 ; Sequence 1453, Application US/09880107
 ; Patent No. US20020142981A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Horne, Darci T.
 ; APPLICANT: Vockley, Joseph G.
 ; APPLICANT: Scherf, Uwe
 ; APPLICANT: Gene Logic, Inc.
 ; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
 ; FILE REFERENCES: 44921-5028-WO
 ; CURRENT APPLICATION NUMBER: US/09/880,107
 ; CURRENT FILING DATE: 2001-06-14
 ; PRIOR APPLICATION NUMBER: US 60/211,379
 ; PRIOR FILING DATE: 2000-06-14
 ; PRIOR APPLICATION NUMBER: US 60/237,054
 ; PRIOR FILING DATE: 2000-10-02
 ; NUMBER OF SEQ ID NOS: 3950
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 1453
 ; LENGTH: 394
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA599526
 US-09-880-107-1453

Query Match	10.1%	Score 392.4	DB 10	Length 394
Best Local Similarity	99.7%	Pred. No. 6.4e-96		
Matches 393	Conservative 0	Mismatches 1	Indels 0	Gaps 0
QY	1709	GATGATCCATCTGCGTTGGCTCCCA	CAGTGTGGATTTACAGGGGTGAGCCACCATGCC	1768
Db	394	GATGATCCATCTGCGTTGGCTCCCA	CAGTGTGGATTTACAGGGGTGAGCCACCATGCC	335

```
Db 154 TTATTTTAAATAAGAAAAACATTCCTAGAGATGATAATTTGTGAAACCTCTTTTGGC 95
QY 2009 TTATTTGCTTTTCCAGATTTAGTCTCTTCTTCCCAATCCGGGAAAGATGTTGAAGAC 2068
Db 94 TTATTTGCTTTTCCAGATTTTATGCTCTCTTCTCCGCATCCGGGAAAGATGTTGAAGAC 35
QY 2069 ATAGGCTAAATTTCTCCAGCCTCACAAATGGTCTT 2102
Db 34 ATAGGCTAAATTTCTCCAGCCTCACAAATGGTCTT 1

RESULT 11
US-09-918-995-19090
; Sequence 19090, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19090
; LENGTH: 344
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(344)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-19090

Query Match 7.5%; Score 289.4; DB 11; Length 344;
Best Local Similarity 98.0%; Pred. No. 66-68;
Matches 293; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1699 CTTGACTTCAGATGATCCATCTGCTTGCCTCCACAGTGTGGATTACAGCGTGAG 1758
Db 41 CTTGACTTCAGATGATCCATCTGCTTGCCTCCACAGTGTGGATTACAGCGTGAG 100
QY 1759 CCACCATGCCCGGCTCTTCTCACCTTTACACCTGTCTTATCTCTCATCTGTTTT 1818
Db 101 CCACCATGCCCGGCTCTTCTCACCTTTACACCTGTCTTATCTCTCATCTGTTTT 160
QY 1819 CACACCTTCATCCCTGTCTTCTCATGTTCACACTGTCTTCCCCCATGTTTCATAGTGC 1878
Db 161 CACACCTTCATCCCTGTCTTCTCATGTTCACACTGTCTTCCCCCATGTTTCATAGTGC 220
QY 1879 TTCTTACATTTGGTTTGAAGGAGCTCTTCTGCTGCTTTTGTGTTTTCCTCCAG 1938
Db 221 TTCTTACATTTGGTTTGAAGGAGCTCTTCTGCTGCTTTTGTGTTTTCCTCCAG 280
QY 1939 AAAATCAGTATTATTTTAAATAAGAAAAACATTCCTAGAGATGATAATTTGTGAAAA 1997
Db 281 AAAATCAGTATTATTTTAAATAAGAAAAACATTCCTAGAGATGATAATTTGTGAAAA 339

RESULT 12
US-09-954-456-1305/c
; Sequence 1305, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
```

```
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1305
; LENGTH: 274
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1305

Query Match 6.8%; Score 263.4; DB 10; Length 274;
Best Local Similarity 97.8%; Pred. No. 61e-61;
Matches 267; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 3599 CTTGGGCTCTCTGGGAAATGGGAAATCGAACAGGACTGAGAGGAGTTACCTTAAGG 3658
Db 274 CTTGGGCTCTCTGGGAAATGGGAAATCGAACAGGACTGAGAGGAGTTACCTTAAGG 215
QY 3659 TCAGAGAAAAACCAAGAGAGTGTGTCTTCTGGAAGCTGAGCTTTCTTTATTCAACTCAT 3718
Db 214 TCAGAGAAAAACCAAGAGAGTGTGTCTTCTGGAAGCTGAGCTTTCTTTATTCAACTCAT 155
QY 3719 TCCCTTCTCCAAATAAGCCACTTGTAGTTGGGCCCTTCCAGGTTGAAGCAAGAGA 3778
Db 154 TCCCTTCTCCAAATAAGCCACTTGTAGTTGGGCCCTTCCAGGTTGAAGCAAGAGA 95
QY 3779 GAAGGCACAGCGTTTGGGAAACAGACATTTCTCGCAATAGCTTCTGCGAAGGAATAAAG 3838
Db 94 GAAGGCACAGCGTTTGGGAAACAGACATTTCTCGCAATAGCTTCTGCGAAGGAATAAAG 35
QY 3839 GATAGAGTGTAAAAATAAAAAAATAAAAAA 3871
Db 34 GATAGAGTGTGGGTTTAAAAAATAAAAAA 2

RESULT 13
US-09-815-343-683/c
; Sequence 683, Application US/09915343
; Patent No. US20010055596A1
; GENERAL INFORMATION:
; APPLICANT: Meagher, Madeleine
; APPLICANT: Xu, Jiangchun E.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.504
; CURRENT APPLICATION NUMBER: US/09/815,343
; CURRENT FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 683
; LENGTH: 279
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-815-343-683
```

Query Match 6.7%; Score 258.6; DB 9; Length 279;
Best Local Similarity 97.8%; Pred. No. 1.2e-59;
Matches 273; Conservative 0; Mismatches 4; Indels 2; Gaps 1;
QY 1681 CCAGGCTGGTCTCGAACTCTGACTTCAGATGATCCATCTGCCTTGGCCCTCCACAGTGC 1740
Db 279 CCAGGCTGGTCTCGAACTCTGACTTCAGATGATCCATCTGCCTTGGCCCTCCACAGAGC 220
QY 1741 TGGATTACAGGCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1800
Db 219 TGGATTACAGGCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 160
QY 1801 TATCCTCACATCTGTTTTCACACCTT--CATCCCTGCTCTCCATGTTTACACACTGTCT 1858
Db 159 TATCCTCACATCTGTTTTCACACCTTTCATCCCTGCTCTCCATGTTTACACACTGTCT 100
QY 1859 TCCCATGTTTCATAGTGCCTTCTTACCAATTTGGTTTGAAGGCGAGTCTTCTCTGGCT 1918
Db 99 TCCCATGTTTCATAGTGCCTTCTTACCAATTTGGTTTGAAGGCGAGTCTTCTCTGGCT 40
QY 1919 TGTCTTTTCTGTTTTCAGAAATCAGTATTATTTT 1957
Db 39 TGTCTTTTCTGTTTTCAGAAATCAGTATTATTTT 1

RESULT 14
US-09-764-891-5534/c
; Sequence 5534, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 5534
; LENGTH: 6670
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-5534

Query Match 6.2%; Score 239.2; DB 11; Length 6670;
Best Local Similarity 86.0%; Pred. No. 1.0e-53;
Matches 265; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
QY 1467 ACACCTATCTTCTCACCTTTTTTTTGGATGAGTCTCGCTCTCTTCCAGGCTGGAG 1526
Db 4664 AGACTTTTCCAGCTTTTTTTTTTTTGAACGAGTCTCACTCTGTCAACCCAGGCTGGAG 4605
QY 1527 TGCAATGGCAGCTTCTCAGCTCACTGCAACCTCCGCTCTTGGTTCAAGCAATTTCTGT 1586
Db 4604 TACAGTGGCAGATCTCAGCTCACTGCAACCTCCACCTCTGTTTCAAGCAATTTCTCC 4545
QY 1587 GCATCAGCTCCCGAGTACTTGGATTACAGGATGCGCAGCAGCAGCAGCAGCAGCAGC 1646
Db 4544 GCCTCAGCTCCCGAGTACTTGGATTACAGGATGCGCAGCAGCAGCAGCAGCAGCAGC 4485
QY 1647 TATTTTATAGACAGCGGTTTTTCCATGTTGGCAGGCTGCTCGACTCTTCACTT 1706
Db 4484 TATTTTATAGACAGCGGTTTTTCCATGTTGGTCAAGGCTGCTCGAACTCCCGACCT 4425
QY 1707 CAGATGATCCATCTGCCTTGGCTCCCAACAGTGTGGGATTACAGGCGTGAGCCACATG 1766
Db 4424 CAGGTGATCCCGGCTTCCGCTCCCAAGTGTGGGATTACAGGCGTGAGCCACATGCG 4365
QY 1767 CCGGCT 1774
Db 4364 CCGGCT 4357

RESULT 15
US-09-764-891-5533/c
; Sequence 5533, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 5533
; LENGTH: 7713
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-5533

Query Match 6.2%; Score 239.2; DB 11; Length 7713;
Best Local Similarity 86.0%; Pred. No. 2e-53;
Matches 265; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
QY 1467 ACACCTATCTTCTCACCTTTTTTTTGGATGAGTCTCGCTCTCTTCCAGGCTGGAG 1526
Db 5708 AGACTTTTCCAGCTTTTTTTTTTTTGAACGAGTCTCACTCTGTCAACCCAGGCTGGAG 5649
QY 1527 TGCAATGGCAGCTTCTCAGCTCACTGCAACCTCCGCTCTTGGTTCAAGCAATTTCTGT 1586
Db 5648 TACAGTGGCAGATCTCAGCTCACTGCAACCTCCACCTCTGTTTCAAGCAATTTCTCC 5589
QY 1587 GCATCAGCTCCCGAGTACTTGGATTACAGGATGCGCAGCAGCAGCAGCAGCAGCAGC 1646
Db 5588 GCCTCAGCTCCCGAGTACTTGGATTACAGGATGCGCAGCAGCAGCAGCAGCAGCAGC 5529
QY 1647 TATTTTATAGACAGCGGTTTTTCCATGTTGGCAGGCTGCTCGAACTCTTGAAT 1706
Db 5528 TATTTTATAGACAGCGGTTTTTCCATGTTGGTCAAGCTGCTCGAACTCCCGACCT 5469
QY 1707 CAGATGATCCATCTGCCTTGGCTCCCAACAGTGTGGGATTACAGGCGTGAGCCACATG 1766
Db 5468 CAGGTGATCCCGGCTTCCGCTCCCAAGTGTGGGATTACAGGCGTGAGCCACATGCG 5409
QY 1767 CCGGCT 1774
Db 5408 CCGGCT 5401

Search completed: January 30, 2004, 19:07:27
Job time : 1311.96 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 30, 2004, 05:43:24 ; Search time 471 Seconds

(without alignments)

3102.903 Million cell updates/sec

Title: US-09-729-674-2

Perfect score: 2130

Sequence: 1 MFCRRGAALLLCVACA.....DDEGEVVEYDDLELEETS 401

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2434939 seqs, 1822278265 residues

Total number of hits satisfying chosen parameters: 4869878

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USFO_spool/US09729674/runat_29012004_095315_12934/app_query.fasta_1.583
-DB=Published Applications NA -OFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFM=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09729674 @CGN 1 1 387 @runat_29012004_095315_12934
-NCPU=6 -ICPU=3 -NO_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=10 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US10D_NEW_PUB.seq.*
- 18: /cgn2_6/ptodata/2/pubpna/US10E_NEW_PUB.seq.*
- 19: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
- 20: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
- 21: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
- 22: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
- 23: /cgn2_6/ptodata/2/pubpna/US10J_PUBCOMB.seq.*
- 24: /cgn2_6/ptodata/2/pubpna/US10K_PUBCOMB.seq.*
- 25: /cgn2_6/ptodata/2/pubpna/US10L_PUBCOMB.seq.*
- 26: /cgn2_6/ptodata/2/pubpna/US10M_PUBCOMB.seq.*
- 27: /cgn2_6/ptodata/2/pubpna/US10N_PUBCOMB.seq.*
- 28: /cgn2_6/ptodata/2/pubpna/US10O_PUBCOMB.seq.*
- 29: /cgn2_6/ptodata/2/pubpna/US10P_PUBCOMB.seq.*
- 30: /cgn2_6/ptodata/2/pubpna/US10Q_PUBCOMB.seq.*
- 31: /cgn2_6/ptodata/2/pubpna/US10R_PUBCOMB.seq.*
- 32: /cgn2_6/ptodata/2/pubpna/US10S_PUBCOMB.seq.*
- 33: /cgn2_6/ptodata/2/pubpna/US10T_PUBCOMB.seq.*
- 34: /cgn2_6/ptodata/2/pubpna/US10U_PUBCOMB.seq.*
- 35: /cgn2_6/ptodata/2/pubpna/US10V_PUBCOMB.seq.*
- 36: /cgn2_6/ptodata/2/pubpna/US10W_PUBCOMB.seq.*
- 37: /cgn2_6/ptodata/2/pubpna/US10X_PUBCOMB.seq.*
- 38: /cgn2_6/ptodata/2/pubpna/US10Y_PUBCOMB.seq.*
- 39: /cgn2_6/ptodata/2/pubpna/US10Z_PUBCOMB.seq.*
- 40: /cgn2_6/ptodata/2/pubpna/US11_PUBCOMB.seq.*
- 41: /cgn2_6/ptodata/2/pubpna/US11_PUBCOMB.seq.*
- 42: /cgn2_6/ptodata/2/pubpna/US11_PUBCOMB.seq.*
- 43: /cgn2_6/ptodata/2/pubpna/US11_PUBCOMB.seq.*
- 44: /cgn2_6/ptodata/2/pubpna/US11_PUBCOMB.seq.*
- 45: /cgn2_6/ptodata/2/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2130	100%	401	1	MFCRRGAALLLCVACA.....DDEGEVVEYDDLELEETS

1	2128	99.9	3871	9	US-09-729-674-1	Sequence 1, Appli
2	1827.5	85.8	1810	12	US-10-108-260A-2058	Sequence 2058, Ap
3	1089.5	51.2	2443	13	US-10-094-749-724	Sequence 724, App
4	1088.5	51.1	2347	10	US-09-962-832-113	Sequence 113, App
5	1088.5	51.1	2347	12	US-10-439-388-64	Sequence 64, Appli
6	567	26.6	2600	14	US-10-045-815-3	Sequence 3, Appli
7	539	25.3	2829	14	US-10-045-815-1	Sequence 1, Appli
8	538	25.3	2322	14	US-10-045-815-5	Sequence 7, Appli
9	538	25.3	2416	14	US-10-045-815-5	Sequence 5, Appli
10	527	24.7	2753	10	US-09-728-952-28	Sequence 28, Appli
11	527	24.7	2753	10	US-09-728-952-51	Sequence 51, Appli
12	498	23.4	3396	13	US-10-094-749-921	Sequence 921, App
13	486	21.4	2865	14	US-10-071-786-86	Sequence 86, Appli
14	388	18.2	398	12	US-10-094-886-23	Sequence 23, Appli
15	321.5	15.1	2305	12	US-10-094-886-23	Sequence 114858,
16	270	12.7	564	13	US-10-027-632-114858	Sequence 114858,
17	270	12.7	564	14	US-10-027-632-114858	Sequence 161474,
18	265.5	12.5	892	13	US-10-027-632-161474	Sequence 161475,
19	265.5	12.5	892	13	US-10-027-632-161475	Sequence 161476,
20	265.5	12.5	892	13	US-10-027-632-161476	Sequence 161476,
21	265.5	12.5	892	14	US-10-027-632-161474	Sequence 161475,
22	265.5	12.5	892	14	US-10-027-632-161475	Sequence 161476,
23	265.5	12.5	892	14	US-10-027-632-161476	Sequence 9365, Ap
24	243	11.4	226	12	US-10-242-535A-9365	Sequence 395, App
25	200.5	9.4	349	9	US-09-925-302-395	Sequence 215, App
26	195	9.2	618	10	US-09-879-536-215	Sequence 1775, Ap
27	168	7.9	150	12	US-10-242-535A-1775	Sequence 2358, Ap
28	161.5	7.6	549	13	US-10-029-386-2358	Sequence 16058, A
29	156	7.3	181	13	US-10-029-386-16058	Sequence 4919, Ap
30	121	5.7	247	12	US-10-242-535A-4919	Sequence 2416, Ap
31	121	5.7	1584	10	US-09-974-300-2416	Sequence 12155, A
32	119	5.6	342	12	US-10-242-535A-12155	Sequence 47173, A
33	105	4.9	1239	12	US-10-359-493-47173	Sequence 1, Appli
34	101	4.7	1230025	12	US-10-289-782-1	Sequence 98, Appli
35	100	4.7	5457	13	US-10-161-051-98	Sequence 53, Appli
36	99.5	4.7	3899	13	US-10-360-522-53	Sequence 6860, Ap
37	99	4.6	2238	13	US-10-032-585-6860	Sequence 266, App
38	99	4.6	4643	10	US-09-864-864-266	Sequence 1504, Ap
39	98.5	4.6	2301	12	US-10-104-047-1504	Sequence 28565, A
40	98	4.6	65	13	US-09-908-975-28565	Sequence 9848, Ap
41	97.5	4.6	2285	15	US-10-198-846-9848	Sequence 117, App
42	96.5	4.5	2249	11	US-09-946-374-117	Sequence 117, App
43	96.5	4.5	2249	12	US-10-011-795A-117	Sequence 117, App
44	96.5	4.5	2249	12	US-10-012-231A-117	Sequence 117, App
45	96.5	4.5	2249	13	US-10-015-387A-117	Sequence 117, App

ALIGNMENTS

RESULT 1

US-09-729-674-1
; Sequence 1, Application US/09729674
; Patent No. US20010039335A1

GENERAL INFORMATION:

- ; APPLICANT: Jacobs, Kenneth
- ; APPLICANT: McCoy, John M.
- ; APPLICANT: LaVallie, Edward R.
- ; APPLICANT: Collins-Racie, Lisa A.
- ; APPLICANT: Evans, Cheryl
- ; APPLICANT: Merberg, David
- ; APPLICANT: Treacy, Maurice
- ; APPLICANT: Agostino, Michael J.
- ; APPLICANT: Steinginger II, Robert J.
- ; APPLICANT: Spaulding, Vikki
- ; APPLICANT: Wong, Gordon G.
- ; APPLICANT: Clark, Hilary
- ; APPLICANT: Fecthel, Kim
- ; APPLICANT: Genetics Institute, Inc.

; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 6055-64X
; CURRENT APPLICATION NUMBER: US/09/729,674
; CURRENT FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: 09/539,330


```
Db 837 GAGGCCAATTGACCCCAATGTTGGTGGCTACTTCTGGGACAAAGTTCTGGCCACCATTG 896
Qy 298 TyrHisTyrLeuGlnPheAlaTyrTyrLysLeuAsnAspLeuLysAsnAlaAlaProCys 317
Db 897 TACCACCTACCTGCGAGTTGGCTACTATAAGTTGATGATGCGCCAGGCTGCCGCGCAGC 956
Qy 318 AlaValSerTyrLeuLeuPheAspGlnAsnAspLysValMetGlnGlnAsnLeuValTyr 337
Db 957 GCCGCCAGCTACAGTCTTCCGACCCCAAGACAGCGCTCATGACAGACAACTGGTGTAT 1016
Qy 338 TyrGlnTyrHisArgAspThrTyrGlyLeuSerAspGluHisPheGlnProCysProGlu 357
Db 1017 TACCGGTTCCACCGGCTCGCTGGGCTGGAAGAGAGAGACTTCCAGCCCGGAGGAG 1076
Qy 358 AlaValGlnPhePheAsnValThrLeuGlnLysGluLeuTyrAspPheAlaLysGlu 377
Db 1077 GCCATGCTCTACCAACACAGACCGCGAGCTGCGGAGCTGCTGGAGTTCAACCCATG 1136
Qy 378 AsnIleMetAspAspGluGlyGluValGluTyrValAspAspLeuLeuGluLeu 397
Db 1137 TACTGCGAGTCAGATGATGAG-----ATGGAGCTG 1166

RESULT 5
US-10-439-388-64
; Sequence 64, Application US/10439388
; Publication No. US2003028617A1
; GENERAL INFORMATION:
; APPLICANT: Aune, Thomas M
; APPLICANT: Olsen, Nancy J
; TITLE OF INVENTION: Method for Predicting Autoimmune Disease
; FILE REFERENCE: 1242/68
; CURRENT APPLICATION NUMBER: US/10/439,388
; CURRENT FILING DATE: 2003-05-16
; PRIOR APPLICATION NUMBER: US 60/381,055
; PRIOR FILING DATE: 2002-05-16
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 64
; LENGTH: 2347
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-439-388-64

Alignment Scores:
Pred. No.: 5,78e-128 Length: 2347
Score: 1088.50 Matches: 216
Percent Similarity: 68.24% Conservative: 59
Best Local Similarity: 53.60% Mismatches: 105
Query Match: 51.10% Indels: 24
DB: 12 Gaps: 5

US-09-729-674-2 (1-401) x US-10-439-388-64 (1-2347)
Qy 4 GlyArgArgGlyAlaAlaAlaLeuLeuAlaLeuLeuCysValAlaCysAlaLeuAlaGala 23
Db 19 GCGTGGCGTGGGGC-----TGCTGTGTTGCTG-CTGGGCGAGC 56
Qy 24 GlyArgAlaGlnTyrGluArgTyrSerPheArgSerPheProArgAspGluLeuMetPro 43
Db 57 GCCGGGCGCAGTACGAGAGAGTACAGCTTCCGGGGCTTCCCGCCCGAGGACCTGATGCCG 116
Qy 44 LeuGluSerAlaTyrArgHisAlaLeuAspLysTyrSerGlyGluHisTyrPalaGluSer 63
Db 117 CTGGCCCGCGGTACGGCGCAGCTCTGGCAGGTACGAGGAGAGAGCTGCGCCGAGAGC 176
Qy 64 ***GlyTyrLeuGluIleSerLeuArgLeuHisArgLeuLeuArgAspSerGluAlaPhe 83
Db 177 GCGCGCTACCTGGAGGGCGGCTCGCGCTGCACCGGCTCTCGCGCAGCAGCGAGGCGCTTC 236
```

RESULT 6

US-10-045-815-3
; Sequence 3, Application US/10045815
; Publication No. US20020160498A1

```
Qy 84 CysHisArgAsnCysSerAla-----AlaProGlnProGluProAlaAlaGlyLeu 100
Db 237 TGCCAGCGCAACTGTCAGCGGCGCCCGCGCGGCAAGCCGATCCCGAGCGCGCGCGC 296
Qy 101 AlaSer-----TyrProGluLeuArgLeuPheGlyLeuLeuArgAlaHisCys 118
Db 297 GCAGACGAGTGGGCGCTCGGAGCTGCGGCTCTTCGGCGCGCTCCTGGAGCGAGCCGCTGC 356
Qy 119 LeuLysArgCysLysGlnGlyLeuProAlaPheArgGlnSerGlnProSerArgGluVal 138
Db 357 CTGGGCGGTGCAAGCGGACGCTGCCGCTTCCAGGTGCCCCCTACCCCGCGGCGAGCTG 416
Qy 139 LeuAlaAspPheGlnArgGluProTyrLysPheLeuGlnPheAlaTyrPheLysAla 158
Db 417 CTGCGTGACTTCCAGAGCGCGCTGCTACCACTACCTACCACTACCGCTGTTCAGGCT 476
Qy 159 AsnAsnLeuProLysAlaIleAlaAlaHisThrPheLeuLeuLysHisProAspAsp 178
Db 477 AACCGCTGGAGAGCGGTGGCGCGCTTACACTTCTCCAGAGAACCCGAGAC 536
Qy 179 GluMetMetLysArgAsnMetAlaTyrTyrLysSer---LeuProGlyValaGluAspTyr 197
Db 537 GAGCTGACCGCCAGTAGTATCTCACTACTATCAGGGGATGCTGGACGTGCGCGACGATCC 596
Qy 198 IleLysAspLeuGluThrLysSerTyrGluSerLeuPheIleArgAlaValArgAlaTyr 217
Db 597 CTCACGACCTAGAGCGCGCGCTTACGAGCGCGTTCCTCCCGGCTGTGAAGCTCTAC 656
Qy 218 AsnGlyGluAsnTrpArgThrSerIleThrAspMetGluLeuAlaLeuProAspPhe 237
Db 657 AACAGCGGGGATTTCCGCGAGCAGCAGCGGAGGACATGGAGCGGCGCTTCTCAGAGTACCTG 716
Qy 238 LysAlaPheTyrGluCysLeuAlaAlaCysGluGlySerArgGluIleLysAspPheLys 257
Db 717 GCAGTCTTTGCCCGGTGCTGCGCGCTGTGAAGGGCGCCATGAGCAGGTGACTTCAAG 776
Qy 258 AspPheTyrLeuSerIleAlaAspHisTyrValGluValLeuGluCysLysIleGlnCys 277
Db 777 GACTTCTACCGGCGCATAGCAGATCTCTTTCAGAGTCCCTTCAGTGCAGGTGAGTGT 836
Qy 278 GluGluAsnLeuThrProValIleGlyTyrProValGluLysPheValAlaThrMet 297
Db 837 GAGGCCAATTTGACCCCAATGTGGTGGCTACTTCGTGGACAAAGTTCTGCGCCACCATG 896
Qy 298 TyrHisTyrLeuGlnPheAlaTyrTyrLysLeuAsnAspLeuLysAsnAlaProCys 317
Db 897 TACCACCTACCTGCGAGTTTGGCTTACTATAGTTGATGATGCGCGAGCTGCCCGCAGC 956
Qy 318 AlaValSerTyrLeuLeuPheAspGlnAsnAspLysValMetGlnGlnAsnLeuValTyr 337
Db 957 GCCCGCAGCTACATGCTCTTCGACCCCAAGGACAGCGCTCATGACAGAGAACCTGGTGTAT 1016
Qy 338 TyrGlnTyrHisArgAspThrTyrGlyLeuSerAspGluHisPheGlnProArgProGlu 357
Db 1017 TACCGGTTCCACCGGCGCTCGCTGGGCGCTGGAAGAGGAGGACTTCCAGCCCGGAGGAG 1076
Qy 358 AlaValGlnPhePheAsnValThrThrLeuGlnLysGluLeuTyrAspPheAlaLysGlu 377
Db 1077 GCCATGCTCTACCAACACAGCGCGCGAGCTGCGGAGCTGCTGGAGTTCACCCACATG 1136
Qy 378 AsnIleMetAspAspGluGlyGluValGluTyrValAspAspLeuLeuGluLeu 397
Db 1137 TACTGCGAGTCAGATGATGAG-----ATGGAGCTG 1166
Qy 398 GluGluThr 400
Db 1167 GAGGAGACA 1175
```



```
Db 234 TCGCGCACGCTGTGCCACCGAAGTCTGCGCGCACCGGACCTGGATCTCGTCCGGAC 293
Qy 96 ProAla-----AlaGlyLeuAlaSerTyrProGluLeuArgLeuPheGlyGly 111
Db 294 CCCAGCCTGAGCAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 353
Qy 112 LeuLeuArgAlaHisCysLeuLysArgCysLysGlnGlyLeuProAlaPheArgGln 131
Db 354 GTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 410
Qy 132 SerGlnProSerArgGluValLeuAlaPheGlnArgArgGluProTyrLysPheLeu 151
Db 411 CTG-----AGTGAAGAACTGGAGTTCACAGCGGAGCGCGCGCGCGCGCGCGCG 464
Qy 152 GlnPheAlaTyrPheLysAlaHisCysLeuLysArgCysLysGlnGlyLeuProAlaPheArgGln 171
Db 465 CAGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 524
Qy 172 LeuLeuLysHisProAspAspGluMetLysArgAsnMetAlaTyrTyrLysSerLeu 191
Db 525 TTGTGGGCAATCCGAGACATGGAGATGCGGAGAACCTGACTATTACCAACCACTG 584
Qy 192 ProGlyAlaGluAspTyr-----ileLysAspLeuGluThrLysSerTyrGluSerLeuPhe 210
Db 585 TCTGGGTGAGGAGCAGACCTCAGGATCTCGAGGCGCAAGCCCATGATGATGATTT 644
Qy 211 IleArgAlaValArgAlaTyrAsnGlyGluAsnTyrPheGlnSerIleThrAspMetGlu 230
Db 645 CGGCTGGGGTACGACTCTTACTCAGAGAGAGAGCCAGGAGAGTGTGCCCGACCTGGAG 704
Qy 231 LeuAlaLeuProAspPhePheLysAlaPheTyrGluCysLeuAlaAlaCysGlyGlySer 250
Db 705 GCGGCACTGACAGAGTACTTTGTGCCCATGAGAGTGCCTGCTCGAGAGGGGCC 764
Qy 251 ArgGluLeuLysAspPheLys-----AspPheTyrLeuSerIleAla 264
Db 765 TATGACTACGCGGTGACAACTACCTAGACTACAGCGCTGACCTCTCCAGGCGCATCACA 824
Qy 265 AspHisTyrValGluValLeuGluCysLysIleGlnCysGluGluAsnLeu---ThrPro 283
Db 825 GATCATTTAGCTCAGGCTCTCACTGTAAAGCAAACTGTGTCAGGAGCTGGCTTCCAC 884
Qy 284 ValIleGlyGlyTyrProValGluLysPheValAlaThrMetTyrHisTyrLeuGlnPhe 303
Db 885 CCAAGTAGGGAAGAACCCCTTTGAAGACTTCTCCCTTCCACTATATTAATCTACAGTT 944
Qy 304 AlaTyrTyrLysLeuAsnAspLeuLysAsnAlaAlaProCysAlaValSerTyrLeuLeu 323
Db 945 GCCTACTACAACATTGGGAAGCTATATACAAAGCTATTGAATGTGCCAAGACCTACTCTC 1004
Qy 324 PheAspGlnAsnAspLysValMetGlnGlnAsnLeuValTyrGlnTyrHisArgAsp 343
Db 1005 TTCTTTCCCATGATGAGGTGATGACACAGAACTGTGGCTTATTAC----- 1049
Qy 344 ThrTrpGlyLeuSerAspGluHis-----PheGlnProArgProGluAlaValGln 360
Db 1050 ACAGCCATGCTTGGAGAAAGAGAGCGCCAGCTCCATCAGCCCGCAGGAGATGCGGAGAA 1109
Qy 361 PhePheAsnValThrThrLeuGlnLysGluLeu-----TyrAspPheAlaLys 376
Db 1110 TACCAGCTGCCAAACCTGTGGAGAAAGAACTGCTTTCTTCGCTTATGATATTTTGA 1169
Qy 377 GluAsnIleMetAspAspGluGluValValGluTyrValAspAspLeuGlu 396
Db 1170 ATTCCCTTTGTGATCCCGGATTCATGGACTCCAGAGAAAGTGAATCCCAAGAGATTGCAA 1229
Qy 397 LeuGluGluThrSer 401
Db 1230 GAGAAACAGAAAGTCT 1244
```

RESULT 9

US-10-045-815-5

```
; Sequence 5, Application US/10045815
; Publication No. US20020160498A1
; GENERAL INFORMATION:
; APPLICANT: Wadawa, Renu
; APPLICANT: Sugihara, Takashi
; APPLICANT: Oribe, Akiko
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; FILE REFERENCES: 06501-091001
; CURRENT APPLICATION NUMBER: US/10/045,815
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: PCT/JP00/02731
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: JP 11/118806
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 2416
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (12)...(2252)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2376_feature
; OTHER INFORMATION: n = A,T,C or G
; US-10-045-815-5
Alignment Scores:
Pred. No.: 1,386-57 Length: 2416
Score: 538.00 Matches: 147
Percent Similarity: 48.71% Conservative: 60
Best Local Similarity: 34.59% Mismatches: 173
Query Match: 25.26% Indels: 46
DB: 13 Gaps: 13
US-09-729-674-2 (1-401) x US-10-045-815-5 (1-2416)
```

```
Qy 5 ArgArgGlyAlaAlaAlaLeuAlaLeuLeuCysValAlaCysAlaLeuArgAlaGly 24
Db 22 CGAAGAGAGGCTGCTGGCAGATGCTAGCGGTGCGCGCGCGCGCGCGCGCGCGCG 81
Qy 25 ArgAlaGlnTyrGluArgTyrSerPheArgSerPheArgAsp---GluLeuMetPro 43
Db 82 -----GCGAGTCTGAGCGGGATGGGACGTGGCAGCC 114
Qy 44 LeuGluSerAlaTyrArgHisAlaLeuAspLysTyrSerGlyGluHisTrpAlaGluSer 63
Db 115 CTGACCTG-CTTTACGACAGGGGACCGCGGCTACTCGCGCAGGAGTGGCCCGGGTG 173
Qy 64 **GlyTyrLeuGluIleSerLeuArgLeuHisArgLeuLeuArgAspSerGluAlaPhe 83
Db 174 GTCTGAACATGAGCGGGCTGTGCTCGCGGGGGCCCTCGGTGCCCTCGCTCGCGC 233
Qy 84 CysHisArgAsnCysSerAla-----AlaProGln-----ProGlu 95
Db 234 TGCGGCACACGCTGTGCCACCGAATCGCGTGGGCACCGGACCTGGATCTCGGTCCGGAC 293
Qy 96 ProAla-----AlaGlyLeuAlaSerTyrProGluLeuArgLeuPheGlyGly 111
Db 294 CCCAGCCTGAGCAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 353
Qy 112 LeuLeuArgAlaHisCysLeuLysArgCysLysGlnGlyLeuProAlaPheArgGln 131
Db 354 GTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 410
Qy 132 SerGlnProSerArgGluValLeuAlaPheGlnArgArgGluProTyrLysPheLeu 151
Db 411 CTG-----AGTGAAGAACTGGAGTTCACAGCGGAGCGCGCGCGCGCGCGCGCG 464
Qy 152 GlnPheAlaTyrPheLysAlaHisCysLeuLysArgCysLysGlnGlyLeuProAlaPheArgGln 171
```


[illegible]

RESULT 10

```

US-09-728-952-28
; Sequence 28, Application US/09728952
; Patent No. US20020111302A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Wang, Jian-rui
; APPLICANT: Wang, Duanrui
; APPLICANT: Yamazaki, Vicki
; APPLICANT: Ujwal, Manusha L.
; APPLICANT: Dmanac, Radje T.
; TITLE OF INVENTION: No. US20020111302A1el Nucleic Acids and
; TITLE OF INVENTION: Polyptides
; FILE REFERENCE: 799
; CURRENT APPLICATION NUMBER: US/09/728,952

```

```
QY 186 AlaTyrTyrLysSerLeuProGluAlaGluAspTyr---IleLysAspLeuGluThrLys 204
Db 760 GACTATTACCAAAACCATCTCTGGAGTGAAGGCGCGACTTCAAGGATCTTGAGACTCAA 819
QY 205 SerTyrGluSerLeuPheIleArgAlaValArgAlaTyrAsnGlyGluAsnTyrArgThr 224
Db 820 CCCCATATGCAAGAAATTTGCACTGGAGTGGGACTTACTCAGAGGAACACCCACAGAA 879
QY 225 SerIleThrAspMetGluLeuAlaLeuProAspPheLysAlaPheTyrGluCysLeu 244
Db 880 GCTGTGCCCCACCTAGAGCGCGCTGCAAGAACTTTTGGCTATGAGGAGTGGCT 939
QY 245 AlaAlaCysGluGlySerArgGluLeuLysAspPheLys-----Asp 258
Db 940 GCCCTCTGCAAGGCGCCCTATGACTACATGGCTACAACTACCTTGTAGTACAACGCTGAC 999
QY 259 PheTyrLeuSerIleAlaAspHisTyrValGluValLeuGluCysLysIleGluCysGlu 278
Db 1000 CTCCTCCAGGCGCATCAGAGATCATTCATCCAGGCTCTCAACTGTAGCAGAACTGTGTC 1059
QY 279 GluAsnLeu---ThrProValIleGlyTyrProValGluLysPheValAlaThrMet 297
Db 1060 ACGGAGCTTGTCTCCCAAGTCGAGAGAGCCCTTTGAAGACTTCTCCCTCCCATCGCAT 1119
QY 298 TyrHisTyrLeuGlnPheAlaTyrTyrLysLeuAsnAspLeuLysAsnAlaProCys 317
Db 1120 TATAATTATCTGCACTTTGCCCTACTATTAACATTTGGGAATTATACACAGGCTGTTGAATGT 1179
QY 318 AlaValSerTyrLeuPheAspGlnAsnAspLysValMetGlnGlnAsnLeuValTyr 337
Db 1180 GCCAAGACCTATCTTCTCTCTCCCAATGACGAGGATGATGAACCAAAATTTGGCCTAT 1239
QY 338 TyrGlnTyrHisArgAspThrTyrGlyLeuSerAspGluHisPheGln-----Pro 354
Db 1240 TATGAGCTATG-----CTTGGAGGAAGAACACACAGCATCCATCGGCGCC 1284
QY 355 ArgProGluAlaValGlnPhePheAsnValThrLeuGlnLysGluLeuTyrAspPhe 374
Db 1285 CGTGAGAGTGCCAAAGAGTACCGACAGCGAGCTACTTGGAAAGAACTCTTTCTTCTC 1344
QY 375 AlaLysGlu-----AsnIleMetAspAspGluGlyGluValValGluTyr 390
Db 1345 GCTTATGATGTTTTTGAATTCCTTTCTGGATCCGGATTCATGACTCCAGAGAAGTG 1404
QY 391 ValAspAspLeuGluLeuGluGluThrSer 401
Db 1405 ATTCACAGAGATTGCAGAGAAACAGAAATCA 1437
```

RESULT 11

```
US-09-728-952-51
; Sequence 51, Application US/09728952
; Patent No. US2002011302A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Kyle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wang, Dunrui
; APPLICANT: Yamazaki, Vicki
; APPLICANT: Ujwal, Manusha L.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US2002011302A1el Nucleic Acids and
; FILE REFERENCE: 799
; CURRENT APPLICATION NUMBER: US/09/728,952
; CURRENT FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 51
; LENGTH: 2753
```

```
TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2418)
; US-09-728-952-51

Alignment Scores:
Pred. No.: 4,35e-56 Length: 2753
Score: 527.00 Matches: 145
Percent Similarity: 43.38% Conservative: 68
Best local Similarity: 29.53% Mismatches: 188
Query Match: 24.74% Indels: 110
DB: 10 Gaps: 12

US-09-729-674-2 (1-401) x US-09-728-952-51 (1-2753)
QY 6 ArgGlyValAlaAlaLeuLeuAlaLeuCysValAlaCysAlaLeuArgAlaGlyArg 25
Db 10 CGCGCCTTGAAGCTGCTGACCACTGCTGGCTGTCTGGCGCGCT-----GCCTCCCAA 63
QY 26 AlaGlnTyrGluArgTyrSerPheArgSerPheProA:AspGluLeuMetProLeuGlu 45
Db 64 GCGAGGTCGAGTCCGAGGACGATGGGCGATGGTGAAGCGCTGATCTGCTC----- 114
QY 46 SerAlaTyrArgHisAlaLeuAspLysTyrSerGlyGluHisTyrAlaGluSer***Gly 55
Db 115 -----TTCCGCGAGGGGACCCGACCTACGCGCGCGGGGACTGGCCCGGGGTGCTCTG 168
QY 66 TyrLeuGluIleSerLeuArgLeuHisArgLeuLeuArgAspSerGluAlaPheCysHis 85
Db 169 AGCATGAACGGCGCTGCTGCCCGGCGAGCCCTCCGCGCGCTTCCGCTGCGCTGCCGC 228
QY 86 ArgAsnCysSerAla-----AlaProGlnProGlu 95
Db 229 ACCCAGTGTGCGCGCGACTTCCCGTGGAGCTGACCCGCGACTGTGCTCCCGCGCGCC 288
QY 96 ProAlaAlaGlyLeuAlaSerTyrProGluLeuArgLeuPheGlyGlyLeuLeuArg 115
Db 289 CAGGCTCTGGGCGCGCGCGCTGCGGACCTGAGCTTCTCGGGGCGCTTCTGCGGTGCG 348
QY 116 AlaHisCysLeuLysArgCysLysGlnGlyLeuProAlaPheArgGlnSerGlnProSer 135
Db 349 GCTGCTGCTGCGCGCTGCTGCTGCGCGCGCGCGCGCGCGCTGCTGCTGCTGCTGCTG 399
QY 136 ArgGluValLeuAlaAspPheGlnArgGluProTyrLysPheLeuGlnPheAlaTyr 155
Db 400 GAAGATGGAGCTGGAGTTCCGCAAGCGGAGCGCCCTACACTACTGCGAGTCCGCTAC 459
QY 156 PheLysAla----- 158
Db 460 TTCAAGGTGCAGACCTGCTCGAAACCGAGCGCGCGGGTCTTCTGGGAGAGAGTGT 519
QY 158 ----- 158
Db 520 GCAGGGACCTGAGGAGCTTGGGGGATCGGGAAGTGTCCGCGAGGGGAAAGTGGCC 579
QY 158 ----- 158
Db 580 TCCTGCTGGGAGCTCTCTCTCGAGCGCGGAGAGCTGCTCTCTCGGAGAGAGCTTCC 639
QY 159 -----AsnAsnLeuProLysAlaIle 165
Db 640 TCGCCCATTCGCATGGCGAGATGTAAACCCAAAGATCAACAAGTGTGAGAAAGCTGT 699
QY 166 AlaAlaAlaHisThrPheLeuLeuLysHisProAspAspGluMetMetLysArgAsnMet 185
Db 700 GCTGAGCAGACACACTTCTTCTGCGCAATCTCTGAGCACATGGAATGAGCAGAACCTA 759
QY 186 AlaTyrTyrLysSerLeuProGlyAlaGluAspTyr---IleLysAspLeuGluThrLys 204
Db 760 GACTATTACCAAAACCATGTCTGGATGAAGGAGCGGCGACTTCAAGATCTTGAGACTCAA 819
```


Db 972 AGAATAGTATTTAGGCTATAGGCTGCTGTATGAAGCTATTCAGATCTACTACAT 1031
Qy 268 IGLUValLeuGluCysLysIleGlnCysGluGluAsnLeuThrProValIleGlyGlyTy 288
Db 1032 CGAGTGCTTGTTCAGCATGAATGTGTGAGGAACTTCCACCCCGCTGCGCGCT 1091
Qy 288 x---ProValGluLysPheValAlaThrMetTyrHisTyrLeuGlnPheAlaTyrTy 307
Db 1092 CTCCTCCACGAGAAATTTCTCTCTGCACTATGATACCTACAGTTTGCTACTATCG 1151
Qy 307 sLeuAsnAspLeuLysAsnAlaProCysAlaValSerTyrLeuLeuPheAspGlnAs 327
Db 1152 AGTGTGAGTATGTGAAGCCCTCGAGTGTCCAAAGCCCTATCTTATCCCATCCAGA 1211
Qy 327 nAspLysValMetGlnGlnAsnLeuValTyrGlnTyrHisArgAspThrTrpGlyLe 347
Db 1212 TGATGAGGATGTCTTACCAATGTGGATTACTATGAG-----AGTCT 1253
Qy 347 userAspGluHisPheGlnPro-----ArgProGluAlaValGlnPhePh 362
Db 1254 GCTGATGATAGCATGTGACCGGCGATCCATTGAGCCAGAGAGGATTTAAACAATGTTGT 1313
Qy 362 eAsnValThrThrLeuGlnLysGluLeuTyrAspPheAlaLysGluAsnIle 379
Db 1314 GAAAGTCTAAGCTGGAGTCTGAGCTGATAAATCAGCTCAGAGAGTCTG 1365

RESULT 13

US-10-071-766-86
; Sequence 86, Application US/10071766
; Publication No. US20020192678A1
; GENERAL INFORMATION:
; APPLICANT: Huei-Mei Chen
; TITLE OF INVENTION: GENES EXPRESSED IN SENESENCE
; FILE REFERENCE: PA-0043 US
; CURRENT APPLICATION NUMBER: US/10/071,766
; CURRENT FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: PERL Program
; SEQ ID NO 86
; LENGTH: 2665
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020192678A1 978730.4
; NAME/KEY: unsure
; LOCATION: 353-387, 694, 727
; OTHER INFORMATION: a, t, c, g, or other
US-10-071-766-86

Alignment Scores:
Pred. No.: 4,898-47 Length: 2665
Score: 456.00 Matches: 123
Percent Similarity: 47.26% Conservative: 58
Best Local Similarity: 32.11% Mismatches: 143
Query Match: 21.41% Indels: 59
DB: 14 Gaps: 11

US-09-729-674-2 (1-401) x US-10-071-766-86 (1-2665)

Qy 19 CysAlaLeuArgAlaGlyArgAlaGlnTyrGluArgTyrSerPheArgSerPheProArg 38
Db 470 TGCTCTAGCCGCGCGCGCGCGCTAC-----499
Qy 39 AspGluLeuMetProLeuGluSerAlaTyrArgHisAlaLeuAspLysTyrSerGlyGlu 58
Db 500 -----TACAGCGGAGAC 511
Qy 59 HisTrpAlaGluSer***GlyTyrLeuGluLeuSerLeuArgLeuHisArgLeuLeuArg 78
Db 512 ---TACGAGCGAGCGGCGCGGAGCTTGAAGCGCGCTGCGCAGCCACCGCGCTCGG 568

Qy 79 AspSerGluAlaPheCysHisArg-AsnCysSerAla-----AlaProG1 93
Db 569 GAAATCCGACGCGCTGTCCCTGCGCTGCGCGCGCCACCGCTCTCCCGCCCCCGC 628
Qy 93 nProGluProAlaAlaGlyLeuAlaSerTyrProGluLeuArgLeuPheGlyGlyLeu 113
Db 629 CCCCCCGCGCGCGCGCGCGCTGAGCTGCCCC-----CTTTCCGCTCCTGTGT 679
Qy 113 uArgArgAlaHisCysLeuLysArgCys-----LysGlnGlyLeuProAlaPheAr 130
Db 680 GGGCGCGCGCGCTNTTATCGCAGCTGTGAGACCCGCGCTCGGGGNGCCCGCATCCCG 739
Qy 130 gGlnSerGlnProSerArgGluValLeuAlaAspPheGlnArgArgGluProTyrLysPh 150
Db 740 CCAC---CCGCTGCGGAGGATGTGCGCAGCACTTCCAGCGCAGAGTGCCTCAACTA 796
Qy 150 eLeuGlnPheAlaTyrPheLysAlaAsnAsnLeuProLysAlaIleAlaAlaHisPh 170
Db 797 CTGCGAGCGGCTACATCAAGCTTAACAGCTCGAAAAGAGCAGTGGAGCGAGCTCACAC 856
Qy 170 rPheLeuLeuLysHisProAspAspGluMetMetLysArgAsnMetAlaTyrTyrLysSe 190
Db 857 ATTTTCGTGGCTAACCTGAGCAGCATGGAATGCGAGCAACATTGAGAATTACAGGC 916
Qy 190 rLeuProGlyAlaGluAsp---TyrIleLysAspLeuGluThrLysSerTyrGluSer 209
Db 917 GACAGCTGTGTGAAGCATTCGAGTTGTAGAGAGAGAGAGCCAGCCACCATCGAGAG 976
Qy 209 uPheIleArgAlaValArgAlaTyrAsnGlyAsnTrpArgThrSerIleThrAspMe 229
Db 977 TTACAATGAGGAGTTAAACATTATGAGCTGATGAGCTTGTGAGATGGCTATCAGCAC 1036
Qy 229 tGluLeuAlaLeuProAspPhePheLysAlaPheTyrGluCysLeuAlaAlaCysGluGl 249
Db 1037 CGAACCAAGCCTTAAGAGAAATTTCTGTTGAAGATACAGATGCCGAGCCCTATGTGAGG 1096
Qy 249 ySerArgGluIleLysAspPheLys-----AspPheTyrLeuSerIle 263
Db 1097 GCTCAGAGATTGAAGATATGAGTATTAGGGTATAAGGCTGCTGTATGAAGCTAT 1156
Qy 263 eAlaAspHisTyrValGluValLeuGluCysLysIleGlnCysGluGluAsnLeuThrPr 283
Db 1157 TGCAGATCACTACATGCGAGTCTTGTGTCAGCATGAATGTGTGAGGAGAACTTGCAC 1216
Qy 283 oValIleGlyGlyTyr---ProValGluLysPheValAlaThrMetTyrHisTyrLeuGl 302
Db 1217 CCGCCCTGCGCGCTCTCTCCCATCGAATTTCTCTCTGCACTATGATTACTACA 1276
Qy 302 nPheAlaTyrTyrLysLeuAsnAspLeuLysAsnAlaAlaProCysAlaValSerTyrLe 322
Db 1277 GTTGCCTACTATCGAGTTGTGAGTATGTGAAGCCCTCGGAGTGTGCGAAAGCCTATCT 1336
Qy 322 uLeuPheAspGlnAsnAspLysValMetGlnGlnAsnLeuValTyrTyrGlnTyrHisAr 342
Db 1337 TCTATGCCATCCAGATGATGAGGATGCTCTGACAAATGTGAGTACTACTATGAG----- 1388
Qy 342 gAspThrTrpGlyLeuSerAspGluHisPheGlnPro-----ArgProG1 357
Db 1389 -----AGTCTGCTGATAGCATTCGACCCGCGCATCCATTCAGGCGCAGAGAGA 1438
Qy 357 uAlaValGlnPhePheAsnValThrLeuGlnLysGluLeuTyrAspPheAlaLysGl 377
Db 1439 TTAAACAATGTTTGAACACCGTCATPAGCTGGAGTCTGAGCTGATAAATCAGCTGCGA 1498
Qy 377 uAsnIle 379
Db 1499 AGGTCTG 1505

RESULT 14

US-10-242-535A-8221
; Sequence 8221, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:

```
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 8221
; LENGTH: 398
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2)..(2)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (263)..(263)
; OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-8221

Alignment Scores:
Pred. No.: 1,04e-39 Length: 398
Score: 388.00 Matches: 81
Percent Similarity: 72.57% Conservative: 1
Best Local Similarity: 71.88% Mismatches: 1
Query Match: 18.22% Indels: 30
DB: 12 Gaps: 1

US-09-729-674-2 (1-401) x US-10-242-535A-8221 (1-398)

Qy 291 GlulysPheValAlaThrMetTyrHisTyrLeuGlnPheAlaTyrTyrLysLeuAsnAsp 310
Db 61 GAGAAATTGGCTACCATGATATATTACTGTGAGTTGCCTATTATTAAGTTGAACGAC 120

Qy 311 LeulysAsnAlaAlaProCysAlaValSerTyrLeuLeuPheAspGlnAsnAspLysVal 330
Db 121 CTGAAGATGACAGCCCTGTGAGTCAGCTATCTGCTTTGATCAGATGACAGGTC 180

Qy 331 MetGlnGlnAsnLeuValTyrGlnTyrHisArgAspThrTrpGlyLeuSerAspGlu 350
Db 181 ATGACGACGAGAACCTGGTGTATTACAGTACCACAGGACACGTTGGGCGCTCTCAGATGAG 240

Qy 351 HisPheGlnProArgProGluAlaValGlnPhePheAsnValThrThrLeuGlnLysGlu 370
Db 241 CACTTCAGCCACACCT----- 258

Qy 371 LeuTyrAspPheAlaLysGluAsnIleMetAspAspGluGly-GluValValGlnTyr 390
Db 259 -----GGAGNAAGTTGTGAATA 276

Qy 390 rVal-AspAspLeuLeuGluLeuGluGlnThrSer 401
Db 277 TGTGGGATGACCTCTTGGAACTGGAGGAGACCCAGC 311

RESULT 15
US-10-094-886-23
; Sequence 23, Application US/10094886
; Publication No: US20040002120A1
; GENERAL INFORMATION:
; APPLICANT: Kameda, Ramesh
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Liu, Xiaohong
; APPLICANT: Spytek, Kimberly A.
```

```
; APPLICANT: Patturajan, Meera
; APPLICANT: Burgess, Catherine
; APPLICANT: Vernet, Corine A.
; APPLICANT: Li, Li
; APPLICANT: Gorman, Linda
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Boldog, Ferenc
; APPLICANT: Guo, Xiaojia
; APPLICANT: Shenoy, Suresh
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Taupier, Raymond J., Jr.
; APPLICANT: Miller, Charles
; APPLICANT: Casman, Stacie
; APPLICANT: Pena, Carol
; APPLICANT: Gangolli, Esha
; APPLICANT: Gusev, Vladimir
; APPLICANT: Smithson, Glennnda
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Gerlach, Valerie
; APPLICANT: Pochart, Pascal
; APPLICANT: Fernandes, Elma
; APPLICANT: Shinkens, Richard
; APPLICANT: Rastelli, Luca
; APPLICANT: Spaderna, Steven
; APPLICANT: LaRochelle, William
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
; FILE REFERENCE: 21402-290 B
; CURRENT APPLICATION NUMBER: US/10/094,886
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/274,322
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/313,182
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/288,052
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: 60/318,510
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/314,018
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: 60/274,194
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/274,849
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/296,693
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: 60/313,626
; PRIOR FILING DATE: 2001-08-21
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 298
; SOFTWARE: Patent in 2.1
; SEQ ID NO 23
; LENGTH: 2305
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (36)..(2244)
US-10-094-886-23

Alignment Scores:
Pred. No.: 5,96e-30 Length: 2305
Score: 321.50 Matches: 103
Percent Similarity: 43.67% Conservative: 59
Best Local Similarity: 27.76% Mismatches: 156
Query Match: 15.09% Indels: 53
DB: 12 Gaps: 12

US-09-729-674-2 (1-401) x US-10-094-886-23 (1-2305)

Qy 48 TyrArgHisAlaLeuAspLysTyrSerGlyClnHisTrpAlaGluSer**GlyTyrLeu 67
```

```
Db      1197 AGGCAGCTCTAC---TATGCCATGGAGCACCTG 1226

Search completed: January 30, 2004, 08:05:27
Job time : 500 secs

156 TACGCTGACGGCTGCGCGCTACGCGCGGGGCTTGGCGCGCGGCTGGCGCTCTG 215
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      68 GlulleSerLeuArgLeuHisArgLeuLeuArgAspSerGluAlaPheCysHisArgAsn 87
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      216 CGGAGGCGCTGCGGAGCAGCGCGGCTGGCGCGGGTGGCGTGGATTGCGGGGCGAGC 275
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      88 CysSerAlaAlaPro-----GlnProGluProAla 97
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      276 TGGCGGCGGATCGGGCGCGCGCTCCCGCGCTTCTCGGGGCGCGGAGCCGAC 335
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      98 Alagly-----LeuAlaSerTyrProGluLeuArgLeuPheGlyGlyLeuLeu 113
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      336 TCCGGGCGGAGCCACACGCGGGTCTGG---GAGCGACAGCTTCTCGGTGCACGCTC 392
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      114 ArgArgAlaHisCysLeuArgCysLysGln-----GlyLeuPro 127
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      393 CGCGCGGAGCTGCTCACCAGTGCACGACGAGGCTGGCCCGGGGCGCGCG 452
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      128 AlaPheArgGlnSerGlnProSerArgGluValLeuAlaAspPheGlnArgGluPro 147
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      453 CGGCTTCGCGTGGGAGCGCGCTCGGGAGCGC-----TTCCGCGCTCGGGAGCCC 503
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      148 TyrLysPheLeuGlnPheAlaTyrPheLysAlaAsnLeuProLysAlaAlaAla 167
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      504 TACAACCTACCTGACAGGCGCTTATCCAGTTGAAGAGCTGGATCTGGCAGCTGCGGCA 563
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      168 AlaHisThrPheLeuLeuLysHisProAspAspGluMetMetLysArgAsnMetAlaTyr 187
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      564 GCACACACCTTCTTTGTAGCAAAACCCATGCACCTGCAGATCGCGGAGGACATGGCTAAG 623
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      188 TyrLysSerLeuProGlyAlaGlu---AspTyrIleLysAspLeuGluThrLysSerTyr 206
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      624 TACAGCAAGATGTCGGGAGTTCGGCCCCAGAGCTTCCGGGACCTGGAGACGCCCCACAC 683
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      207 GluSerLeuPheIleArgAlaValArgAlaTyrAsnGlyGluAsnTyrArgThrSerIle 226
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      684 TGGCGAGCTATGACACTGGCTGGAGCTACTGGGGCGCCAGGAGGAGGAGTGGGCTGCGAGCCAG 743
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      227 ThrAspMetGluLeuAlaLeuProAspPhePheLysAlaPheTyrGluCysLeuAlaAla 246
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      744 CCCAGGCTAGAGAGGAGCTCTTCAGGGGAGCCTGCGCCAGATGGAGAGCTGCCGTGCTGAC 803
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      247 CysGluGlySerArgGluIleLysAspPheLysAsp----- 258
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      804 TGTGAGGGGCTGAGAGCAGCAGCGGGGCTGAGAGAGGAGGAGTGGGGCTGCGAGCCAG 863
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      259 -----PheTyrLeuSerIleAlaAspHisTyrValGluValLeuGluCysLysIleGln 276
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      864 GGGGGCCTCTATGAGGCCATTGCAGGACACTGGATTTCAGTCTGCACTGCCGCGCAACGC 923
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      277 Cys---GluGluAsnLeuThrProValIleGlyTyrProValGluLysPheValAla 295
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      924 TGTGTGGGGAGAGCCACAGCCCTGGTGGCAGCTTCCCTGTCCTCCAGACTTCTTCCC 983
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      296 ThrMetTyrHisTyrLeuGlnPheAlaTyrTyrLysLeuAsnAspLeuLysAsnAlaAla 315
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      984 AACCAAGCTGAGCGGCTACATGAGGCCCATGCTCAGTGGCAATCTGTCCAGGCTATA 1043
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      316 ProCysAlaValSerTyrLeuLeuPheAspGlnAsnAspLysValMetGlnGlnAsnLeu 335
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1044 GAAATGTCCTGAGTGCTCTCTTACCCGGAGGATGAGGCTGCCAGAGGGGCTCTG 1103
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      336 ValTyrTyrGlnTyrHisArgAspThrTrpGlyLeuSerAspGluHisPheGlnProArg 355
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1104 AACCAGTACAGGCCCGCTGGGA-----GAGCCGAGA 1136
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      356 Pro-----GluAlaValGlnPhePheAsnValThrThrLeu-----Gln 368
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1137 CCTGGCCTCGGAGAGGAGACATCCAGCGCTTTCATCCCGCATCCCTCGGGGAGAG 1196
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      369 LysGluLeuTyrAspPheAlaLysGluAsnIle 379
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 30, 2004, 05:18:19 ; Search time 81 Seconds

(without alignments)

2185.119 Million cell updates/sec

Title: US-09-729-674-2

Perfect score: 2130

Sequence: 1 MEPCRRGAALLLVCVACA.....DDEGEVVEYVDDLELEETS 401

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09729674/runat 29012004.095314 12909/app query.fasta_1.583
-DB=Issued Patents.NA -QPWT=fastap -SURFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=BLOSUM62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFM=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09729674 @cgn2_1.56 @runat 29012004.095314 12909 -NCFU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents.NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	195	9.2	618	3	US-09-328-111-215
2	140	6.6	11309	4	US-08-961-527-108
C 3	106	5.0	1664976	4	US-08-916-421B-1
4	102.5	4.8	80161	3	US-09-036-987A-1
5	102.5	4.8	80161	3	US-09-370-700-1
6	102.5	4.8	80161	3	US-09-603-207-1
7	101.5	4.8	1131	4	US-09-252-991A-7006
8	101	4.7	1230025	4	US-09-198-452A-1
9	100.5	4.7	1557	4	US-09-252-991A-7033
C 10	100.5	4.7	2751	4	US-09-252-991A-7066
11	100.5	4.7	5833	4	US-09-221-0175-545
12	99.5	4.7	2316	4	US-09-252-991A-6950

13	97	4.6	963	4	US-09-252-991A-6979	Sequence 6979, Ap
14	97	4.6	2233	4	US-09-252-991A-15346	Sequence 15346, A
C 15	97	4.6	2463	4	US-09-252-991A-15430	Sequence 15430, A
16	97	4.6	2784	4	US-09-252-991A-15252	Sequence 15252, A
17	96.5	4.5	2079	4	US-09-489-847-25	Sequence 25, Appl
C 18	95.5	4.5	1230	4	US-09-252-991A-10343	Sequence 10343, A
19	95.5	4.5	1389	4	US-09-252-991A-10807	Sequence 10807, A
20	94.5	4.4	4184	2	US-08-785-310A-4	Sequence 4, Appli
21	94.5	4.4	8948	4	US-09-643-597-119	Sequence 119, App
22	94.5	4.4	8948	4	US-09-480-884A-119	Sequence 119, App
23	94.5	4.4	8948	4	US-09-542-615A-119	Sequence 119, App
24	94.5	4.4	8948	4	US-09-606-421B-119	Sequence 119, App
25	94	4.4	32758	4	US-09-604-694B-1	Sequence 1, Appli
26	94	4.4	34303	2	US-08-735-609-4	Sequence 4, Appli
27	94	4.4	34303	2	US-08-735-609-4	Sequence 4, Appli
28	94	4.4	34303	3	US-09-315-372-4	Sequence 4, Appli
29	94	4.4	34303	3	US-09-244-752-4	Sequence 4, Appli
30	94	4.4	34303	3	US-09-245-497-4	Sequence 4, Appli
31	94	4.4	34303	4	US-09-562-919-4	Sequence 6, Appli
32	94	4.4	34382	2	US-08-374-483-6	Sequence 3, Appli
33	94	4.4	35408	3	US-08-973-334-3	Sequence 3, Appli
34	94	4.4	35408	3	US-09-563-869A-3	Sequence 3, Appli
35	94	4.4	35408	3	US-08-549-489-3	Sequence 3, Appli
36	94	4.4	35935	2	US-08-735-609-1	Sequence 1, Appli
37	94	4.4	35935	2	US-08-735-609-1	Sequence 1, Appli
38	94	4.4	35935	3	US-08-379-452-43	Sequence 43, Appli
39	94	4.4	35935	3	US-09-315-372-1	Sequence 1, Appli
40	94	4.4	35935	3	US-09-244-752-1	Sequence 1, Appli
41	94	4.4	35935	3	US-09-245-497-1	Sequence 1, Appli
42	94	4.4	35935	3	US-09-409-670-43	Sequence 43, Appli
43	94	4.4	35935	4	US-09-562-919-1	Sequence 1, Appli
44	93	4.4	651	4	US-09-252-991A-11750	Sequence 11750, A
45	93	4.4	765	4	US-09-252-991A-11829	Sequence 11829, A

ALIGNMENTS

RESULT 1
US-09-328-111-215/c
; Sequence 215, Application US/09328111
; Patent No. 6262333
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astie, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Derti, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/328,111
; CURRENT FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: US 60/088,801
; EARLIER FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 215
; LENGTH: 618
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(618)
; OTHER INFORMATION: n = A,T,C or G
US-09-328-111-215

Alignment Scores:
 Pred. No.: 2.17e-13 Length: 618
 Score: 195.00 Matches: 45
 Percent Similarity: 80.00% Conservative: 3
 Best Local Similarity: 75.00% Mismatches: 9
 Query Match: 9.15% Indels: 3
 DB: 3 Gaps: 0

US-09-729-674-2 (1-401) x US-09-328-111-215 (1-618)

QY 345 TrpGlyLeuSerAspGluHisPheGlnProArgProGluAlaValGlnPheAsnVal 364
 Db 515 TGGGGCCCTTNGAATGAGCANTCCAGCCAGNCTTAAGACAGTTCAGTTCTTAAATGTG.456
 QY 365 ThrThr-LeuGlnLysGlu-LeuTyxAspPhe-AlaLysGluAsnIleMetAspAsp 383
 Db 455 NCCACCACTCCAGANGAAGCTTGATGACITTTGGNTAAGGAATATAATGGATGAT 396
 QY 384 GluGlyGluValValGluTyxValAspLeuLeuGluGluThrSer 401
 Db 395 GAGGAGAAGTGTGGAATATGTGGATGACCTCTTGGAACTGGAGGAGACCACT 342

RESULT 2

US-08-961-527-108
 ; Sequence 108 Application US/08961527
 ; Patent No. 6420135
 ; GENERAL INFORMATION:
 ; APPLICANT: Charles Kunsch
 ; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
 ; NUMBER OF SEQUENCES: 391
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville
 ; STATE: Maryland
 ; COUNTRY: USA
 ; ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 COMPUTER: HP Vectra 486/33
 OPERATING SYSTEM: MSDOS version 6.2
 SOFTWARE: ASCII Text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/961,527
 FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Brookes, A. Anders
 REGISTRATION NUMBER: 36,373
 REFERENCE/DOCKET NUMBER: PB340P1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504
 TELEFAX: (301) 309-8512
 INFORMATION FOR SEQ ID NO: 108:

SEQUENCE CHARACTERISTICS:

LENGTH: 11309 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 US-08-961-527-108

Alignment Scores:

Pred. No.: 5.82e-05 Length: 11309
 Score: 140.00 Matches: 91
 Percent Similarity: 36.67% Conservative: 70
 Best Local Similarity: 20.73% Mismatches: 152
 Query Match: 6.57% Indels: 126
 DB: 4 Gaps: 22

US-09-729-674-2 (1-401) x US-08-961-527-108 (1-11309)

QY 13 AlaLeuLeuCysValAlaCysAla-----LeuArgAlaGlyArgAlaGlnTyr 28
 Db 4169 GCATACACTGCCAGCTTGTGGAGCCAGCTCCATTTGCTCAAGGACCAAGTGTACGG 4228
 QY 29 GluArgTyrSerPheArgSerPheProArgAspGluLeuMetProLeuGluSerAlaTyr 48
 Db 4229 ACGCATTTTGGCCATAAATCTTAAAGACGTGATTTTTTTTGTGAAATGAAGTCCA 4288
 QY 49 ArgHisAlaLeuAspLysTyr-SerGlyGluHisTrpAlaGluSer**GlyTyrLeuGlu 68
 Db 4289 GAACACCTGGCCAAATPAGGAATCCCTCTATCAGCTGGTTGAAATAAGACAAAAGTTCAA 4348
 QY 69 IleSerLeuArgLeuHisArgLeuLeuArgAspSerGluAlaPheCysHisArgAsn--- 87
 Db 4349 TTAGTAGTACCCGCTTTCAGAACTTAAACAGATTGGGATGTATTTGTAAATGGCAATCTA 4408
 QY 88 -----CysSerAlaAlaProGln-----ProGluProAlaAlaGly 99
 Db 4409 GCTCTAGAAGTTTCAGTGTAGTCCCTTGCTTCAGAAAGTCTTAAAGACGCAAGTGAAGGC 4468
 QY 100 LeuAlaSer-----TyrProGluLeuArgLeuPheGlyGlyLeuLeuArgArgAlaHis 117
 Db 4469 TATCGTAGTCAGGGTTACCAAGTACTGTGGTTGCTGGGTCAAAACTGTGCTCAAGGAG 4528
 QY 118 CysLeuLysArgCysLysGlnGlyLeuProAlaPheArgGlnSer----- 132
 Db 4529 CGTTTGACTCGTCTACAGCAAGGTTTCTTTATTTTTCAGTCAAAACATGGGCTTTATGTT 4588
 QY 133 -----GlnProSerArgGluValLeu-----AlaAspPhe 142
 Db 4589 TGGGAATTAGACAAGGAAAAACAAAGTTTAAAGACTCAAAATACCTGATTACACGATCTC 4648
 QY 143 GlnArgArgGluProTyrLysPheLeuGlnPheAlaTyrPheLysAlaAsn----- 159
 Db 4649 CGCGTAAACTCCATTTATCAATCAAGCAATTTTCTTATGCTCAAGGTAGTTATTGGAA 4708
 QY 160 -----AsnLeuPro-----LysAlaIleAlaAlaHisThrPheLeuLeuLysHisPro 176
 Db 4709 ATATTGCGTCTTCCCTATPAGAGACAAAAATATCTCATTTTACAGTTCTTCGAGACAAG 4768
 QY 177 AspAspGluMetMetLysArgAsnMetAlaTyrTyrLysSerLeuProGlyAlaGluAsp 196
 Db 4769 GACATCTGCTGCTATATCCGCAACAACCTTTATTTATCAAAATCTC-----TTT 4816
 QY 197 TyrIleLysAspLeuGluThrLysSerTyrGluSerLeuPheIleArgAlaValArgAla 216
 Db 4817 TGGATGAAGAA---CAAGCAGAAAGCCTATCAAA----- 4846
 QY 217 TyrAsnGlyGluAsnTrpArgThr-SerIleThrAspMetGluLeuAlaLeuProAspPhe 236
 Db 4847 ---AAGGAGAAATATCTCTGACT----- 4867
 QY 237 PheLysAlaPheTyrGluCysLeuAlaCysGluGlySerArgGluIleLysAspPhe 256
 Db 4868 -----TATGGACTG 4876
 QY 257 LysAspPheTyrLeuSerIleAlaAspHisTyrValGluValLeuGluCysLysIleGln 276
 Db 4877 AAGAATGGTAT----- 4888
 QY 277 CysGluGluAsnLeuThrProValIleGlyGly---TyrProValGluLysPheValAla 295
 Db 4889 -----CCACAAATTCGACCAATAGTGGCAAAATTTTCCAGATTGAACAGACTTGACT 4942
 QY 296 ThrMetTyrHisTyrLeuGlnPheAlaTyrTyrLys-----LeuAsnAspLeuLys 312
 Db 4943 AGCTATTATCAGCACTTT---TATACCTATTACCAAAAAAATCTCTCAAAATGATTGCCAA 4999
 QY 313 AsnAlaAlaProCysAlaVal-----SerTyrLeuLeuPheAsp----- 325
 Db 5000 AAGCTTTATCCACGAGCTTTTATCAGCAATATTTCTTTGAAATAATATGGTAGAATAGAA 5059

NAME/KEY: misc_feature	
LOCATION: {163385}..{163385}	
OTHER INFORMATION: n equals a	t, c, or g
NAME/KEY: misc_feature	
LOCATION: {191989}..{191989}	
OTHER INFORMATION: n equals a	t, c, or g
NAME/KEY: misc_feature	
LOCATION: {191995}..{191995}	
OTHER INFORMATION: n equals a	t, c, or g
NAME/KEY: misc_feature	
LOCATION: {231980}..{231980}	
OTHER INFORMATION: n equals a	t, c, or g
NAME/KEY: misc_feature	
LOCATION: {234187}..{234187}	
OTHER INFORMATION: n equals a	t, c, or g
NAME/KEY: misc_feature	
LOCATION: {234220}..{234220}	
OTHER INFORMATION: n equals a	t, c, or g
NAME/KEY: misc_feature	
LOCATION: {234814}..{234814}	
OTHER INFORMATION: n equals a	t, c, or g
NAME/KEY: misc_feature	
LOCATION: {309398}..{309398}	
OTHER INFORMATION: n equals a	t, c, or g
NAME/KEY: misc_feature	
LOCATION: {309418}..{309418}	
OTHER INFORMATION: n equals a	t, c, or g
NAME/KEY: misc_feature	
LOCATION: {312837}..{312837}	
OTHER INFORMATION: n equals a	t, c, or g
NAME/KEY: misc_feature	
LOCATION: {312993}..{312993}	
OTHER INFORMATION: n equals a	t, c, or g
NAME/KEY: misc_feature	
LOCATION: {319226}..{319226}	
OTHER INFORMATION: n equals a	t, c, or g
NAME/KEY: misc_feature	
LOCATION: {559167}..{559167}	
OTHER INFORMATION: n equals a	t, c, or g
NAME/KEY: misc_feature	
LOCATION: {559241}..{559241}	
OTHER INFORMATION: n equals a	t, c, or g
NAME/KEY: misc_feature	
LOCATION: {600992}..{600992}	
OTHER INFORMATION: n equals a	t, c, or g
NAME/KEY: misc_feature	
LOCATION: {622708}..{622708}	
OTHER INFORMATION: n equals a	t, c, or g
NAME/KEY: misc_feature	
LOCATION: {657081}..{657081}	
OTHER INFORMATION: n equals a	t, c, or g
NAME/KEY: misc_feature	
LOCATION: {657203}..{657203}	
OTHER INFORMATION: n equals a	t, c, or g
NAME/KEY: misc_feature	
LOCATION: {674435}..{674435}	
OTHER INFORMATION: n equals a	t, c, or g
NAME/KEY: misc_feature	
LOCATION: {682442}..{682442}	
OTHER INFORMATION: n equals a	t, c, or g
NAME/KEY: misc_feature	
LOCATION: {713652}..{713652}	
OTHER INFORMATION: n equals a	t, c, or g
NAME/KEY: misc_feature	
LOCATION: {741684}..{741684}	
OTHER INFORMATION: n equals a	t, c, or g
NAME/KEY: misc_feature	
LOCATION: {779455}..{779455}	
OTHER INFORMATION: n equals a	t, c, or g
NAME/KEY: misc_feature	
LOCATION: {779676}..{779676}	
OTHER INFORMATION: n equals a	t, c, or g
NAME/KEY: misc_feature	

LOCATION: (855539) ..(855539)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (871619) ..(871619)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1084830) ..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1096846) ..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1119881) ..(1119881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1130881) ..(1130881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1310988) ..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1313224) ..(1313224)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1349473) ..(1349473)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1349491) ..(1349491)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1470091) ..(1470091)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1569020) ..(1569020)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1602912) ..(1602912)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1603734) ..(1603734)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1637998) ..(1637998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1664855) ..(1664855)
OTHER INFORMATION: n equals a, t, c, or g

US-08-916-421B-1
Alignment Scores:
Pred. No.: 1.58e+03
Score: 106.00
Percent Similarity: 38.49%
Best Local Similarity: 25.10%
Query Match: 4.98%
DB: 4
Length: 1664976
Matches: 60
Conservative: 32
Mismatches: 80
Indels: 68
Gaps: 12

US-09-729-674-2 (1-401) x US-08-916-421B-1 (1-1664976)

QY 181 MetLysArgAsnMetAlaTyrTyrLysSerLeuProGlyAlaGluAspTyrIleLysAsp 200
Db 495629 CTTAAAGAGTGTGGCATTCCAAATGGGATTTAAACT-GCTCCAAAGTATATAAG--- 495574
QY 201 LeuGluThrLysSerTyrGluSerLeuPheIleArgAlaValArgAlaTyrAsnGly--- 219
Db 495573 -----AAAGAGTTTTATCA-----GCCTACTTCGGTTCG 495544
QY 220 -----GluAsnTrpArgThrSerIleThrAspMetGluLeuAla 232
Db 495543 GAATTAACAACCTCCAAAATTAGAAATCATGGGACTTCATTTAAAGAGTATCATTCAAA 495484
QY 233 LeuProAspPhePheLysAlaPheTyrGluCysLeuAlaCysGluGlySerArgGlu 252

Db 495483 ATAGCTAAGATTGAGGAGATATTTGACGAA-----GATAGATTG 495445
QY 253 IleLysAspPheLysAspPheTyrLeuSerIleAlaAspHisTyrValGluValLeuGlu 272
Db 495444 ATTAAGACATTAAGAGATGCTAAAGAAATTTGGT-----ATTGAA 495403
QY 273 CysLysIleGlnCysGluGlu---AsnLeuThrProValIleGlyTyrProValGlu 291
Db 495402 TTAAGGTTAGGTTGAAGAGGCAATCTAAGAAA-----GATGGTTATAAACAA 495349
QY 292 LysPheValAlaThrMetTyrHisTyrLeuGlnPhe-----AlaTyrTyrLys 307
Db 495348 GTTATGTGGTCAATATACACCAATAAAGAAATTTGGAGAAATGGATATACCTAT 495289
QY 308 LeuAsnAspLeuLysAsnAlaAlaProCysAlaValSerTyrLeuLeuPhe----- 324
Db 495288 GCTAACAAAAAGAAACATTAGCAAGATATGCTATGAATACCTCTTACAAAAGAGAAA 495229
QY 325 -----AspGlnAsnAspLysValMetGlnAsnLeuValTyrTyrGlnTyrHis 341
Db 495228 TATCTAAAGACAGAAACATTAAATAAATAACCAAAATTTATACATTTGAT 495169
QY 342 ArgAspThrTrpGlyLeuSerAspGluHisPheGlnProArgProGluAlaValGlnPhe 361
Db 495168 AAG-----TTT 495163
QY 362 PheAsnValThrThrLeuGlnLysGluLeuTyrAspPheAlaLysGluAsnIleMetAsp 381
Db 495162 ATAAAGAGAAATGCCTTAAGAA-----GGATTTGTAAGAAAAAAATTTGAAGT 495112
QY 382 AspAspGluGlyGluValValGluTyrValAspAspLeuLeuGluGluThr 400
Db 495111 ATTGAAGAAACAAA---GTAGATTATGCTATGATATAACTACATCTCAGAAACT 495058

RESULT 4

US-09-036-987A-1
; Sequence 1, Application US/09036987A
; Patent No. 6143526
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H.
; APPLICANT: Broughton, Mary C.
; APPLICANT: Crawford, Kathryn P.
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Merlo, Donald J.
; APPLICANT: Treadway, Patti J.
; APPLICANT: Turner, Jan R.
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dow Agrosciences LLC Patent Department
; STREET: 9330 Zionsville Road
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: USA
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/036,987A
; FILING DATE: 09-MAR-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stuart, Donald R.
; REGISTRATION NUMBER: 28,479
; REFERENCE/DOCKET NUMBER: 50,608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317)337-4816
; TELEFAX: (317)337-4847

```
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80161 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-036-987A-1

Alignment Scores:
Pred. No.: 31.3 Length: 80161
Score: 102.50 Matches: 72
Percent Similarity: 31.64% Conservative: 34
Best Local Similarity: 21.49% Mismatches: 119
Query Match: 4.81% Indels: 110
DB: 3 Gaps: 15

US-09-729-674-2 (1-401) x US-09-036-987A-1 (1-80161)
QY 73 LeuHisArgLeuLeuArgAsp-SerGluAlaPheCysHisArgAsnCysSerAlaAlaPr 92
Db 19414 ATCGAGCGAGCTTTGGCGGACATTGGAGCGGCTGCTCCACAGTTGGCGGCTGCTC 19473
QY 92 oGlnProGlu---ProAlaAlaGlyLeuAla----- 101
Db 19474 GGTGTGGCGCTGACCGAAGCGGATTTGGAAAGCGAGCTGCGTGGTTCACCGGTATC 19533
QY 102 -----SerTyrProGluLeuArgLeuPheGlyGlyLeuLeuArgAlaHisCy 118
Db 19534 CCGGTCCTGCTGCTCCCGCGGATC-----CCGCGCGTGGTGAGACTGCACCATG 19584
QY 118 sLeuLysArgCysLysGlnGlyLeuProAlaPheArgGlnSerGlnPro----- 134
Db 19585 CGCGTCGCTGCAGCTCTGTGGTGGCGGATCAGGACCTCGTCCCGGATGTTGTCGAG 19644
QY 135 -----Se 135
Db 19645 CCGCCAGCAGCGGCGACCATTCGCATGAGATGGCTGGATAGGGAAGGAAACGACGAG 19704
QY 135 rArgGluValLeuAlaAspPheGlnArgGluProTyrLysPheLeuGlnPheAlaTy 155
Db 19705 TACGGCATACTT-----CGGACCCAGTCTCTTTCCCGCGATAGGCGCA 19749
QY 155 r----- 156
Db 19750 GCAGCCCTACTCCATTGGCGGAGATTGGAAATGCGCTGCTATGTCATCGCGTT 19809
QY 156 eLysAlaAsnAsnLeuProLysAlaAlaAlaAlaHisThrPheLeuLeuLysHisPr 176
Db 19810 GACGTCAACAGGACTTCGCGCGGCAACAATAGTGTGTAC----- 19849
QY 176 oAspAspGluMetLysArgAsnMetAlaTyrTyrLysSerLeuProGlyAlaGluAs 196
Db 19850 -----GGCAGGATGTCACGCGACCATCGAAGATCTTTGGTGGCGCGCA 19893
QY 196 pTyrIleLysAspLeuGluThrLysSerTyrGluSerLeuPheIleArgAlaValArgAl 216
Db 19894 CCGGTTTCACGGAACGAGTGAATGCGGAGCTCC-----GC 19932
QY 216 aTyrAsnGlyGluAsnTrpArgThrSerIleThrAspMetGluLeuAlaLeuProAspPh 236
Db 19933 TCGATCGGG-----TGGCGCGGACCTGTACGCTGATCACCGTTGTTGTCGGGGATT 19986
QY 236 ePhe---LysAlaPheTyrGluCysLeu-----Al 245
Db 19987 CATGGGAAGATTTCGCTGCTGCTTTGCTCTCTGCGCGGATAGTTATAGTCGTACCGC 20046
QY 245 aAlaCysGlySerArgGluIleLysAsp-----PheLysAspPheTyrIleuSerIl 263
Db 20047 CGCATCGCGGTAACCGCGAATTAACGAGCGGTAGTTTGGCGTCTTTCTCTCTGTGT 20106
QY 263 eAlaAspHisTyrValGluValLeuGluCysLysIleGlnCysGluGluAsnLeuThrPr 283
Db 263 eAlaAspHisTyrValGluValLeuGluCysLysIleGlnCysGluGluAsnLeuThrPr 283
```

QY 155 r-----Ph 156
Db 19750 GCAGCCCTACTCCATTGGCCAGGATTTGGAAATGCGCTGATGTCGATGCGCGTT 19809
QY 156 elysalaasnleuprolysalalalealalaahisthrpheleuleuLeuLysHisPr 176
Db 19810 GAGCTCCACGACITTCGGCGGCAACAAATAGTGTGTCC-----19849
QY 176 oAspAspGluMetMetLysArgAsnMetAlaTyrTyrLysSerLeuProGlyAlaGluAs 196
Db 19850 -----GGCAGGAATGTACCGGACCATCGAAGATCTTTGGTTCGCCGCA 19893
QY 196 pTyrileLysAspLeuGluThrLysSerTyrGluSerLeuPheLeuArgAlaValArgAl 216
Db 19894 CTGTGTTTACCGCAACGAGTGAATGCGGAGCTCC-----GC 19932
QY 216 aTyrAsnGlyGluAsnTTPArgThrSerIleThrAspMetGluLeuAlaLeuProAspPh 236
Db 19933 TCGATCGGG-----TGGCCGGACCTGTACGGTGATCACCGTTGGTTCTTCGGGGATT 19986
QY 236 ePhe---LysAlaPheTyrGluCysLeu-----Al 245
Db 19987 CATGGGAAGATTTCGCTGGCTGTTTGCCTCTCGCGGATAGTTATAGTCGTCACGC 20046
QY 245 aAlaCysGluGlySerArgGluileLysAsp-----PheLysAspPheTyrLeuSerIl 263
Db 20047 CCATGCGCGGTAAACCGGAATTAACGCGGTAGTTTTCGCGTCTTTCTCTGTGT 20106
QY 263 eAlaAspHisTyrValGluValLeuGluCysIleGlnCysGluGluAsnLeuThrPr 283
Db 20107 GTTCTCTGCTGCTCCAGAAATTAACGAAAGGTGAACGTTGCAGAGATCAGGCATACC 20166
QY 283 o'valle---GlyGlyTyrProValGluLysPheValalaThrMetTyrHisTyrLeuGl 302
Db 20167 GTGTGTCAGGTGCGCACCAATCGCAGCGGTTCGGCAGATGATGACCTG-----20221
QY 302 nPheAlaTyrTyrLysLeuAsnAspLeuLysAsnAla-----AlaProCysAlaVa 319
Db 20222 -----GTCAGCCGCTTGTGACTCGTTCGGCGGCGCCCTCGGCCAT 20265

RESULT 6

US-09-603-207-1
; Sequence 1, Application US/09603207B
; Patent No. 6521406
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H
; APPLICANT: Broughton, Mary C
; APPLICANT: Crawford, Kathryn P
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Treadway, Patti J
; APPLICANT: Turner, Jan R
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; FILE REFERENCE: 50489 DIV1
; CURRENT APPLICATION NUMBER: US/09/603,207B
; EARLIER FILING DATE: 2000-06-23
; EARLIER APPLICATION NUMBER: 09/370,700
; EARLIER FILING DATE: 1998-03-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 80161
; TYPE: DNA
; ORGANISM: Saccharopolyspora spinosa
US-09-603-207-1

Alignment Scores:

Pred. No.: 31.3 Length: 80161

Score: 102.50 Matches: 72
Percent Similarity: 31.64% Conservative: 34
Best Local Similarity: 21.49% Mismatches: 119
Query Match: 4.81% Indels: 110
DB: 4 Gaps: 15
US-09-729-674-2 (1-401) x US-09-603-207-1 (1-80161)
QY 73 LeuHisArgLeuLeuArgAsp-SerGluAlaPheCysHisArgAsnCysSerAlaAlaPr 92
Db 19414 ATCAGCAGAGCTTTGGCGACATTTGACGCGGTCTCTCCACAGTTGCGCGCTGCCCTC 19473
QY 92 oGlnProGlu---ProAlaAlaGlyLeuAla-----101
Db 19474 GTGTTCGCGCTGACCAACCGCGGATTTGGAAAGCGCAGCTCGGTGTTTCCACCGTATC 19533
QY 102 -----SerTyrProGluLeuArgLeuPheGlyGlyLeuLeuArgArgAlaHisCyl 118
Db 19534 GCGGTCTCTGCTGCTCCCGCGGATC-----CCGCGCGGTGGTGCAGACTGCACCATG 19584
QY 118 sLeuLysArgCysLysGlnGlyLeuProAlaPheArgGlnSerGlnPro-----134
Db 19585 CGCGGTGCGCTGCGAGCTCTGTGTGTCGCGCGATCAGACCTCGTCCCGATGCTTGCAG 19644
QY 135 -----Se 135
Db 19645 CGCCAGCACAGCGGCACCATTTGCCATGAGATGCTCGGATAGGCAAGGGAACGACGAG 19704
QY 135 rArgGluValLeuAlaAspPheGlnArgGluProTyrLysPheLeuGlnPheAlaTyr 155
Db 19705 TACGCGCATACTT-----CGGACCCAGTCTCTTCCCGCGATAGCGCA 19749
QY 155 r-----Ph 156
Db 19750 GCAGCCCTACTCCANTGCCAGGATTTGGAATGCGCTCGTATGTCGATGCGCGTT 19809
QY 156 elysAlaAsnAsnLeuProLysAlaAlaAlaAlaAlaHisThrPheLeuLeuLysHisPr 176
Db 19810 GAGCTCCACGACITTCGGCGGCAACAAATAGTGTGTAC-----19849
QY 176 oAspAspGluMetMetLysArgAsnMetAlaTyrTyrLysSerLeuProGlyAlaGluAs 196
Db 19850 -----GGCAGGAATGTACCGGACCATCGAAGATCTTTGGTTCGCCGCA 19893
QY 196 pTyrileLysAspLeuGluThrLysSerTyrGluSerLeuPheLeuArgAlaValArgAl 216
Db 19894 CTGTGTTTACCGGAACGAGTGAATGCGCGAGCTCC-----GC 19932
QY 216 aTyrAsnGlyGluAsnTTPArgThrSerIleThrAspMetGluLeuAlaLeuProAspPh 236
Db 19933 TCGATCGGG-----TGGCGCGGACCTGTACGGTGATCACCGTTGGTTCTTCGGGGATT 19986
QY 236 ePhe---LysAlaPheTyrGluCysLeu-----Al 245
Db 19987 CATGGGAAGATTTCGCGTGTGTTTGCCTCTCGCGGATAGTTATAGTCGTCACGC 20046
QY 245 aAlaCysGluGlySerArgGluileLysAsp-----PheLysAspPheTyrLeuSerIl 263
Db 20047 CGCATCGCGGTAAACCGGAATTAACGCGGTAGTTTTCGCGTCTTTCTCTGTGT 20106
QY 263 eAlaAspHisTyrValGluValLeuGluCysLysIleGlnCysGluGluAsnLeuThrPr 283
Db 20107 GTTCTCTGCTGCTCCAGAAATTAACGAAAGGTGAACGTTGCAGAGATCAGGCATACC 20166
QY 283 o'valle---GlyGlyTyrProValGluLysPheValalaThrMetTyrHisTyrLeuGl 302
Db 20167 GTGTGTCAGGTGCGCACCAATCGCAGCGAGTTGGGCGAGATGATGACCTG-----20221
QY 302 nPheAlaTyrTyrLysLeuAsnAspLeuLysAsnAla-----AlaProCysAlaVa 319
Db 20222 -----GTCAGCCGCTTGTGAACTCGTTCGGCGGCGCCCTCGGCCAT 20265
QY 319 lSerTyrLeuLeuPheAspGlnAsnAspLysValMetGlnGln 333

Db 484 GCGCTTCGTCACAGCAGTACGCTCGACGAGCGCTCGCAACTATCGCGGGCTGCTCT 543

Qy 211 eArgAlaValArgAlaTyrAsnGlyGluAsnTyrArg 223

Db 544 G---GCCGAGCGCGCGCGCGCGCGCATACCTGGCGC 577

RESULT 10

US-09-252-991A-7066/c

; Sequence 7066, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 7066

; LENGTH: 2751

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-7066

Alignment Scores:

Pred. No.:	0.241	Length:	2751
Score:	100.50	Matches:	46
Percent Similarity:	41.62%	Conservative:	26
Best Local Similarity:	26.59%	Mismatches:	63
Query Match:	4.72%	Indels:	38
DB:	4	Gaps:	9

US-09-729-674-2 (1-401) x US-09-252-991A-7066 (1-2751)

Qy 78 ArgAspSerGluAlaPheCysHisArgAsnCysSerAlaAlaProGlnProGluProAla 97

Db 2668 AATCTGTCAGGCGCATTTCCAGCGTCGCGGATAGGGAGCAATCGGCTTCGCTCA 2609

Qy 98 AlaGly-----LeuAlaSerTyrProGluLeuArgPheGlyGlyLeuArg 114

Db 2608 TCGAAGCAGCAAGGATCTTCGATGCGCGCACTCTTCGC-----CCCGTATTCTGCGC 2555

Qy 115 ArgAla-HisCysLeuArgCysLysGlnGlyLeuProAlaPheArgGlnSerGlnPr 134

Db 2554 CGCGCTCTCTGCTG-----TACCAGCGCTGCTCCAGGCTGCTCCAGGCTGCGCGC 2513

Qy 134 oSerArgGluValLeuAlaAspPheGlnArg-----ArgGluProTyr 148

Db 2512 CGAAGCGAAATCTCGCGCTGTTCAGCGCTGGCAGCGCGCCAGCGCGCGCAGCGG 2453

Qy 148 rLysPheLeuGlnPheAlaTyrPhe-----LysAlaAsnLe 161

Db 2452 CAAGCGCTCGTGAAGGGTCTATCGCGCGCTGACTATGCGCGCGCGCGCAGCGG 2393

Qy 161 uProLysAlaIleAlaAlaHisThrPheLeuLeuLysHisProAspAspGluMetMe 181

Db 2392 CGAGCGCACCTCTCGCGCAAGCAGCGCTTCTCGCGCGCCATCCGACTTCGAG----- 2338

Qy 181 tLysArgAsnMetAlaTyrTyrLysSerLeuProGlyAlaGluAsp----- 196

Db 2337 -CAGCGCATGCTCGCTCGCGCGAGATCGTGGCGCGCGCGCGCGCGCGCGCGCGCG 2279

Qy 197 ----TyrIleLys-----AspLeuGluThrLysSerTyrGluSerLeuPhe1 211

Db 2278 GCCTTCGTCACAGCAGTACGCTCGACGAGCGCTGCGCAACTATCCGGGCTGCTCT 2219

Qy 211 eArgAlaValArgAlaTyrAsnGlyGluAsnTyrArg 223

Db 2218 G---GCCGAGCGCGCGCGCGCGCGCATACCTGGCGC 2185

RESULT 11

US-09-221-017B-545

; Sequence 545, Application US/09221017B

; Patent No. 6444799

; GENERAL INFORMATION:

; APPLICANT: Ross, Bruce C.

; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF

; NUMBER OF SEQUENCES: 1120

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MORRISON & FOERSTER

; STREET: 755 PAGE MILL ROAD

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304-1018

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: Windows

; SOFTWARE: FastSeq for Windows Version 2.0b

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/221,017B

; FILING DATE: 23-DEC-1998

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PP1182

; FILING DATE: 31-DEC-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PP1546

; FILING DATE: 30-JAN-1998

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PP2911

; FILING DATE: 09-APR-1998

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/AU98/01023

; FILING DATE: 10-DEC-1998

; ATTORNEY/AGENT INFORMATION:

; NAME: Monroy, Gladys H

; REGISTRATION NUMBER: 32,430

; REFERENCE/DOCKET NUMBER: 27340-20021.00

; TELEPHONE: 650-813-5600

; TELEFAX: 650-494-0792

; TELEX: 706141

; INFORMATION FOR SEQ ID NO: 545:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 5633 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: circular

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: UNKNOWN

; ORIGINAL SOURCE:

; ORGANISM: PORPHYROMONAS GINGIVALIS

; NAME/KEY: misc.feature

; LOCATION: 1...5633

US-09-221-017B-545

Alignment Scores:

Pred. No.:	0.761	Length:	5633
Score:	100.50	Matches:	92
Percent Similarity:	36.20%	Conservative:	68
Best Local Similarity:	20.81%	Mismatches:	165
Query Match:	4.72%	Indels:	118
DB:	4	Gaps:	23

US-09-729-674-2 (1-401) x US-09-221-017B-545 (1-5633)

Qy 31 TyrSerPheArgSerPheProArgAspGluLeuMetProLeuGluSerAla----- 47

Db 31 TyrSerPheArgSerPheProArgAspGluLeuMetProLeuGluSerAla----- 47

Db 1796 TATCTTCTCGGGCGTGGCAGCAGCAATCTGGCAAAATCGATTGAGCATTCAGGAC 1855
Qy TyrArgHisAlaLeuAspLysTyrSerGlyGluHisTrpAlaGluSer***GlyTyrLeu 67
Db 1856 TACCGTCGGGATGGAGCTGATCGCAACGATGAG-----GGCATGCTG 1900
Qy GluLeuSerLeuArgLeuHisArgLeuLeuArgAspSerGluAlaPheCysHis----- 85
Db 1901 -----GTGAATCTGACCGGAGTGTCTACCGACAGCAAAACG-CTATGCCGAAGCAGC 1950
Qy 86 -----ArgAsnCysSerAlaAlaProGlnProGluProAla----- 97
Db 1951 GGAAGGGGTAGCGGACTGCTCAAAATCTACCCCAAAAGCAAGCAACATATAGTCT 2010
Qy 98 -----AlaGlyLeuAlaSerTyrProGluLeuArgLeuPheGlyGlyLeu 113
Db 2011 CAGCGCATAGAGTGGCGAGCAGGATAC----- 2040
Qy 114 ArgArgAlaHisCysLeuLysArgCysLysGlnGly-Leu----- 126
Db 2041 CGTCGACCCAT-----TCGCGAACTGAACGAGGACTTCGGATGGAATCTCTCTTTGC 2094
Qy 127 -ProAlaPheArgGln-----SerGlnProSerArgGluValLe 139
Db 2095 TCCTGCTATTGCGAGATGGCATCTGCACCTCAAGAGCAAGCGCAACGAGGAGGCTAT 2154
Qy 139 uAlaAspPheGlnArgArg-----GluProTyrLysPheLeuGlnPhe----- 153
Db 2155 GGCAGCCCTGCACAAAGCAATAGAAATTCGAGCCGAGCAACTGTCCAACTATATCAACCG 2214
Qy 154 -----AlaTyrPheLysAlaAsnLeuProLysAlaIleAlaAlaHisThrPh 171
Db 2215 CGGTGTCATCGCTACCAATCGAAGACTTGGTGGAGCCATGGACGACTACACCCAGT 2274
Qy 171 eLeuLeuLysHisProAspAspGluMetMetLysArgAsnMetAlaTyrTyrLysSerLe 191
Db 2275 CGTTCGCAAAAGCGCAACGACAGTTGGCACGTTTCAACCGTGCTTGTCCGCTCTTA 2334
Qy 191 uproGly-----AlaGluAspTyr-----IleLysAspLeuGluThrLy 204
Db 2335 CTGCGTACGTGAACAATGCGATCGAGGACTTCGATGATGATCATCCCTCTGGAGCCGGA 2394
Qy 204 SerTyrGluSerLeuPheIleArgAlaValArgAlaTyrAsnGlyGluAsnTrpArgTh 224
Db 2395 GAACATATCATGCACTCTCAACCGTGCCATCTCTCCTCAGCAGATCAACGAGAACCGAAA 2454
Qy 224 rSerIleThrAspMetGluLeuLeuLeu-----ProAspPhePheLysAlaPheTy 241
Db 2455 AGCCATTACCGACTTCGACAAAGTCTGGGACACTATCTCTGACTTGTGCTGGCTACTA 2514
Qy 241 rGluCysLeuAlaAlaCysGluGlySerArgGluIleLysAspPhe-----Ly 257
Db 2515 T-----GTCGCTCCCAAGCCAAAGAGCATTTGGCGACACTCGCGAGCGGAACG 2565
Qy 257 sAspPheTyrLeuSerIle----- 263
Db 2566 GGACTACTGGCGGCTTTTCGACATAGAAAAAGCGGCACAGACAAACAAAAACGAA 2625
Qy 264 ----AlaAspHisTyrValGluValLeuGluCysLysIleGlnCysGluGluAsnLeuTh 282
Db 2626 AAAAGGACGCTTTCGACTTCACCAAGAGACAGCTGAGGATACCGACGAAACG----- 2680
Qy 282 rProValIleGlyGlyTyrProValGluLysPheValAlaThrMetTyrHisTyrLeuGl 302
Db 2681 -----ATAGAGAAGTTC-----AACCTGCT 2700
Qy 302 nPheAlaTyrTyrLysLeuAsnAspLeuLysAsnAlaAlaProCysAlaValSerTyrLe 322
Db 2701 CGTGATATCCGAGAAATCTCCGAGAAAGCGCACAGCTACTCCAGTCGGATCGTGGGCG 2760
Qy 322 uLeuPheAspGlnAsnAspLysValMetGlnGlnAsnLeuValTyrTyrGlnTyrHisAr 342
Db 2761 GATTGAGGACACGATGTAGAGGTATACAGCCCTCGCGGCTATTTCGTCCTCAGCTACTAGA 2820

Qy 342 gaSpThrTrpGlyLeuSerAspGluHisPheGlnProArg-----ProGluAlaVa 359
Db 2821 GCAG-----TCTCTTCGCAAGAGGTGGCGCGTCTACTACTCGAAGCCAT 2868
Qy 359 lGlnPhePheAsn---ValThrThrLeuGlnLysGluLeuTyrAspPheAlaLysGluAs 378
Db 2869 TGCTGCTTCAACGCAAAAGAGTTCTGCCAAGGAGCTCAAAACCGATCAACCGTGAGT 2928
Qy 378 nileMetAspAspAspGluGlyGluValValGluTyrValAspAspLeuLeuGluLeuGl 398
Db 2929 AGCTCTGAGTCAAGAG-----CAGTGGCATATCAGCAACAGGACATACGGAATATC 2982
Qy 398 uGlu 399
Db 2983 CGAG 2986

RESULT 12

US-09-252-991A-6950
; Sequence 6950, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107195.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 6950

; LENGTH: 2316

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-6950

Alignment Scores:

Pred. No.: 0.24 Length: 2316
Score: 99.50 Matches: 92
Percent Similarity: 35.49% Conservative: 56
Best Local Similarity: 22.06% Mismatches: 129
Query Match: 4.67% Indels: 140
Gaps: 25

US-09-729-674-2 (1-401) x US-09-252-991A-6950 (1-2316)

Qy 1 MetGluProGlyArgArg-----GlyAlaAlaAlaLeuAlaLeuLeuCys 16

Db 510 CTGCAACCAAGCGCTCGACAGATCGGCAAGGACGGCATGCT----- 551

Qy 17 ValAlaCysAla-----LeuArgAlaGlyArgAlaGlnTyrGluArgTyrSerPhe 33

Db 552 GTCTCAGCCGCTGATCCCGCATTCAGGCGCTCGCGCAATATCCT-----GAAAT 605

Qy 34 ArgSerPheProArgAspGluLeuMetProLeuGluSerAlaTyrArg-----HisAla 51

Db 606 CGAAGTCTGCACCGCGCGCAT-----GAAACAGCAGTTCGAAGCCTCCGCTCG 653

Qy 52 LeuAspLysTyrSerGlyGluHis-----TrpAlaGluSer***GlyTyrLeuGluLeuSer 70

Db 654 GTCCGCGGATACAGGCGAGTCGCATCCACGCGAAAGTCGTCACACCGCGTGAGATTGA 713

Qy 71 LeuArgLeuHisArgLeuLeuArgAspSerGluAlaPheCysHisArgAsnCysSerAla 90

Db 714 CGTCGATATCACCGA-----CCGATTCCGGTCTATGGCGCATCAACAT----- 758

Qy 91 AlaProGlnProGluProAlaAlaGlyLeuAlaSerTyrProGluLeuArgLeuPheGly 110

Db 759 -----GGTCCAGAAATACAAAGCGGAGTACACCGTACACCGTACCAATGGAAGC 812

QY 111 -----GlyLeuLeuArgAlaHisCysLeuLysArgCysLysGlnGlyLeuProAla 128
Db 813 CAAAGCAATCTCTACATGAACCCGATGTGTGGCGC----- 851
QY 129 PheArgGlnSerClnProSerArgGlu-----ValLeuAlaAspPheGlnArgGluPro 147
Db 852 -----CTCACATGGAAGCAATCTGTGGAGGTTTCAGAGGGAACAGCAA 896
QY 148 Tyr-LysPheLeuGlnPheAlaTyrPheLysAlaAsnLeuProLysAlaIleAlaAl 167
Db 897 CACCAAGGTCATCAAA-----GACCCATCTCGC 926
QY 167 alaHisThrPheLeuLeuLysHisProAspAspGluMet-----MetLysArgAs 184
Db 927 GGGCCACGGAAGCAATCTCACAGATGATCCAGGACATCCGGCCGACAGCCACAGAT 986
QY 184 nMetAlaTyrTyrLysSerLeuProGlyAlaGluAspTyrIleLysAspLeuGluThrLy 204
Db 987 CGTCGCGGGCAAAAGCCGGCTCCGGC-----GTGATCGATGCCGAGCCAA 1034
QY 204 sSerTyrGluSerLeuPheIleArgAlaValArgAlaTyrAsnGlyGluAsnTrpArgTh 224
Db 1035 CAAATACGAGAACTCTCGCGCCCGCTAAAGCTGCGCTCACAACG----- 1086
QY 224 rSerIleThrAspMetGluLeuAlaLeuProAspPhePheLysAlaPheTyrGluCysLe 244
Db 1087 -----AGTATCGGGCTCGAGTCACTTCCTCAAGGCC----- 1119
QY 244 uAlaAlaCysGluGlySerArgGluIleLysAspPheLysAspPheTyrLeuSerIleAl 264
Db 1120 -----GCCAAGGAT----- 1128
QY 264 aAspHisTyrValGluValLeuGluCysLysIleGlnCysGluGluAsnLeuThrProVa 284
Db 1128 ----- 1128
QY 284 lIleGlyGlyTyrProValGluLysPheValAlaThrMetTyrHisTyrLeuGlnPheAl 304
Db 1129 -----GTGTTCGCGGAACCTGGTACACCATC----- 1167
QY 304 aTyrTyrLysLeuAspLeuLys-----AsnAlaAlaProCysAlaValSe 320
Db 1168 -----GTCAAGATGAACGAGCTGTAGTCAGTCCAGCTGAACGGCTTCCGCGGAGAAC 1223
QY 320 rTyrLeuLeuPheAspGlnAsnAspLysValMetGlnGlnAsnLeuValTyrGlnTy 340
Db 1224 CATTCATCTGTCGACAAAGAGACCAAGGAGACATCCAGAC----- 1266
QY 340 rHisArgAspThrTrpGlyLeuSerAspGluHisPheGlnProArgProGluAlaValGl 360
Db 1267 -TATGCTGACGTGATGCTTACCGCTGACGAATATGCAAGCACCGCGTGGACTCAGTT-- 1323
QY 360 nPhePheAsnValThrThrLeuGlnLysGluLeuTyrAspPheAlaLysGluAsnIleMe 380
Db 1324 -----AGGAACACTAC---TATCGGAAACCAACATCG 1355
QY 380 tAspAspAspGluGlyGlu---ValValGluTyrValAspAspLeuLeu 395
Db 1356 AGAGCCCGGTGACGGCGAAGGGCGCTGGGAGTACGTCAAGAGATGATG 1404
RESULT 13
US-09-252-991A-6979
; Sequence 6979, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6979
; LENGTH: 963
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-6979
Alignment Scores:
Pred. No.: 0.115 Length: 963
Score: 97.00 Matches: 39
Best Local Similarity: 44.53% Conservative: 22
Query Match: 28.47% Mismatches: 43
Indels: 33
Caps: 7
US-09-729-674-2 (1-401) x US-09-252-991A-6979 (1-963)
QY 111 GlyLeuLeuArgAlaHisCysLeuLysArgCysLysGlnGlyLeuProAlaPheAr 130
Db 66 GGTATTTCGCGCGCTCTTCGCTG-----TACCCAGCGTGGT 107
QY 130 gGlnSerGlnProSerArgGluValLeuAlaAspPheGlnArg----- 144
Db 108 CCAGGCTGCCGCGAAGCGAATCTCCGCTGTCCAGCGTGGCAGCGCCGCCACGG 167
QY 145 -ArgGluProTyrLysPheLeuGlnPheAlaTyrPhe-----Ly 157
Db 168 GCAGCGCAGCGCGAAGCGCTGGAAGGGTCTATGCGCGCTGGACTACTATGCGCG 227
QY 157 sAlaAsnAsnLeuProLysAlaIleAlaAlaHisThrPheLeuLeuLysHisProAs 177
Db 228 CGCCAGACCGCGGAGCGCACCTCTCCGCCAAGCAGGCGCTTTCGCGCGCATCCGA 287
QY 177 pAspGluMetMethLysArgAsnMetAlaTyrTyrLysSerLeuProGlyAlaGluAsp-- 196
Db 288 CTTTCGAG-----CAGCGCATCGTGTGCGCGCAGATCGTCCGCGGAGCAGGAAG 341
QY 197 -----TyrIleLys-----AspLeuGluThrLysSerTyrGl 207
Db 342 CCGTACGAGTGGCTTCGTCAAGCAGGACGCTCGAGGAGTACGCTCGAGGAGCGCTGGCACTATCC 401
QY 207 uSerLeuPheIleArgAlaValArgAlaTyrAsnGlyGluAsnTrpArg 223
Db 402 GGGCTCTCTCTG---GCCGAGCGCGCGCGCGCGCATACCTGGCGC 447
RESULT 14
US-09-252-991A-15346
; Sequence 15346, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 15346
; LENGTH: 2283
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-15346
Alignment Scores:
Pred. No.: 0.458 Length: 2283
Score: 97.00 Matches: 87
Percent Similarity: 30.24% Conservative: 37

Best Local Similarity: 21.22% Mismatches: 125
Query Match: 4.55% Indels: 161
DB: 4 Gaps: 17

US-09-729-674-2 (1-401) x US-09-252-991A-15346 (1-2283)

QY 22 ArgAlaGlyArgAlaGlnTyrGlu-----ArgTyrSerPheArgSerPhe 36
DB 1087 CGGCTGGAAGACGAGCGCTCCCGGTACTGGTCAAGGACGCTCGCTGGCGCGCAATT 1146
QY 37 ProArgAspGluLeuMetProLeuGluSerAlaTyrArgHisAlaLeuAspLysTyrSer 56
DB 1147 CCCGCTGATGCTGCGTACCTGATGATCCGCGCACCG-----1185
QY 57 GlyGluHisTrpAlaGluSer**GlyTyrLeuGluHisLeuArgLeuHisArgLeu 76
DB 1186 -----CGCGTGTTCGCTTC 1200
QY 77 LeuArgAspSerGluAlaPheCysHisArgAsnCysSerAlaAlaProGlnProGluPro 96
DB 1201 CTTCCGCGCTCA-----CCCGAGCTT 1221
QY 97 AlaAlaGlyLeuAlaSerTyrProGluLeuArgLeuPheGlyGlyLeuLeuArgAla 116
DB 1222 CGAGTGGCCCTGGAGCGCAGCTGACGAAGTCTCCAGGGGGCGACCTTCGAGGCT 1281
QY 117 His-----CysLeuLysArgCysLysGlnGlyLeu 126
DB 1282 GAACGACCTGCCCGACCGACCTTCGAAAGCATCTCGCTGACCGAGCGCAACACTTCGT 1341
QY 127 ProAlaPheArgGlnSerGlnProSerArgGluValLeuAlaAspPheGlnArg-ArgG1 146
DB 1342 CGAGCATTTTCATGATTCCAG---CGGCTGTGTCTTCGCGCTTCTTCAGCGCAAGC 1398
QY 146 uProTyrLysPheLeuGlnPheAlaTyrPheLysAlaAsnLeuProLysAlaIleAl 166
DB 1399 CGATTTCCCTGCTGCGAGTGGACTTCACCGCGGAGGCGGAGTGCACACGCGGAGGA 1458
QY 166 alaAlaHisThrPheLeuLeuLysHisProAspAspGluMetMetLysArgAsn-----184
DB 1459 AGCCCGACCTGTTCCGACCTCTC-----GAACCATGGGCAAGGCTCA 1506
QY 185 -MetAlaTyrTyrLysSerLeu-----ProGly-----193
DB 1507 CATGCGGTGTACGAGACCTCGCGCTACGCGCTCGCGATCTCTGTGCGCGGCTACTC 1566
QY 194 -----AlaGluAspTyrIleLysAspLeuGluThrLysSerTyrGluSerLe 209
DB 1567 GGAATCTATCCGTTGGAAGACCTGATCTGGGACCAACCAACGAGCGCTCGCG-----1621
QY 209 uPheIleArgAlaValArgAlaTyrAsnGlyGluAsnTrpArgThrSerIleThrAspMe 229
DB 1622 -----TTCGCGAGGACATCTCAACCT 1644
QY 229 tGluLeuAlaLeuProAspPhePheLysAlaPheTyrGluCysLeuAlaLysGluG1 249
DB 1645 GCACCGCTCGACGCGCGCGCTCGGGCGTTCGTGGAGCGCTGGAAGACTC-----1699
QY 249 ySerArgGluLysAspPheLysAspPheTyrLeuSerIleAlaAspHisTyr-----267
DB 1700 -----GAGGTGGACGACTACCGACATCACCACGCTGATCGCGCTGAGTTCGACGA 1752
QY 268 -----ValGluValLeuGluCysLysIleGlnCysGluGluAsnLe 281
DB 1753 CAACAGCTTGGGCGCAACTGACCATCTCTCGAGCTGAAGTG-----1795
QY 281 uThrProValIleGlyGlyTyrProValGluLysPheValAlaThrMetTyrHisTyrLe 301
DB 1796 -----CTCATCGC-----1804
QY 301 uGlnPheAlaTyrTyrLysLeuAsnAspLeuLysAsnAlaAlaProCysAlaValSerTy 321
DB 1805 -----CTCGCCCTGAAGCGCTTCGAAGACCGCAAGGAGCGGTGGAGCGTTCCTCCAGTA 1860

QY 321 rLeuLeuPheAspGlnAsnAspLysValMetGlnGlnAsnLeuValTyrGlnTyrHi 341
DB 1861 C-----AAGGACACAGCGCTCGAGCGCGCTGTTCTAC-----1894
QY 341 sArgAspThrTrpGlyLeuSerAspGluHisPheGlnProArgProGluAlaValGlnPh 361
DB 1895 -----CAGGC 1899
QY 361 ePheAsnValThrThrLeuGlnLysGluLeuTyrAspPheAlaLysGluAsnIleMetAs 381
DB 1900 GCTGAACGTGGTCTG-----GAGTGTCTCT 1926
QY 381 pAspAspGluGlyGluValValGluTyr 390
DB 1927 CGACGACGAGCTGGAGATCGCGACTAC 1954

RESULT 15

US-09-252-991A-15430/C
; Sequence 15430, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 15430
; LENGTH: 2463
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15430

Alignment Scores:
Pred. No.: 0.517 Length: 2463
Score: 97.00 Matches: 87
Percent Similarity: 30.24% Conservative: 37
Best Local Similarity: 21.22% Mismatches: 125
Query Match: 4.55% Indels: 161
DB: 4 Gaps: 17

US-09-729-674-2 (1-401) x US-09-252-991A-15430 (1-2463)

QY 22 ArgAlaGlyArgAlaGlnTyrGlu-----ArgTyrSerPheArgSerPhe 36
DB 1204 CGGCTGGAAGACGAGCGCTCCCGTACTGCTCAAGGACGCTCGCTGGCGCGCAATT 1145
QY 37 ProArgAspGluLeuMetProLeuGluSerAlaTyrArgHisAlaLeuAspLysTyrSer 56
DB 1144 CCGGTGATGCTGCTGACCTGATGAATCCGCGCACCG-----1106
QY 57 GlyGluHisTrpAlaGluSer**GlyTyrLeuGluHisLeuArgLeuHisArgLeu 76
DB 1105 -----CGGCTGTTCGCTTC 1091
QY 77 LeuArgAspSerGluAlaPheCysHisArgAsnCysSerAlaAlaProGlnProGluPro 96
DB 1090 CTTCCGCGCTCA-----CCCGAGCTT 1070
QY 97 AlaAlaGlyLeuAlaSerTyrProGluLeuArgLeuPheGlyGlyLeuLeuArgAla 116
DB 1069 CGAGTGGCCCTGGAGCGCAGCTGACGAAGTCTCGAGCGCGCAGCTTCGAGGCT 1010
QY 117 His-----CysLeuLysArgCysLysGlnGlyLeu 126
DB 1009 GAACGACCTGCCCCAGCGACCTTCGAAAGCATCTGCGCTGACCGAGCGCAACACTTCGT 950

```
QY 127 ProAlaPheArgGlnSerGlnProSerArgGluValLeuAlaAspPheGlnArg-ArgGl 146
Db 949 CGAGCATTTTCATCGATTCCAG---CGCGTGGTGTCTCTGGCGCTTCTTCAGCGCCCAAGGC 893
QY 146 uProTyrLysPheLeuGlnPheAlaTyrPheLysAlaAsnAsnLeuProLysAlaIleAl 166
Db 892 CGATTTCCTCCCTTCGTGGAGTGGACTTACCCGCGAGCGGAGGCTGCCAACGCCGAGGA 833
QY 166 aAlaAlaHisThrPheLeuLeuLysHisProAspAspGluMetMetLysArgAsn----- 184
Db 832 AGCCGCGAGCCCTGTTGGCATCTC-----GAAGCCATGGGCAAGCAGGTCTTA 785
QY 185 -MetAlaTyrTyrLysSerLeu-----ProGly----- 193
Db 784 CATGGCGGTGTACGAAACCTCGCGCTACCGGCTCCCGGATCTCTGGTCCCGGCTACTC 725
QY 194 -----AlaGluAspTyrIleLysAspLeuGluThrLysSerTyrGluSerLe 209
Db 724 GGAATCTATCCCGTGGAGACCTGATCTGGGACACACCAACAGGCCCTGGCG----- 670
QY 209 uPheIleArgAlaValArgAlaTyrAsnGlyGluAsnTyrArgThrSerIleThrAspMe 229
Db 669 -----TTCGCGAGGACATCTCAACCT 647
QY 229 tGluLeuAlaLeuProAspPhePheLysAlaPheTyrGluCysLeuAlaAlaCysGluGl 249
Db 646 GCACCGCTCGACGAGCGCGCTCGCGGCTGTGGAGCGCTCGAGACTGC----- 592
QY 249 ySerArgGluIleLysAspPheLysAspPheTyrLeuSerIleAlaAspHisTyr----- 267
Db 591 -----GAGGTGGAGCTATACCGGACATCACCGCTGATCGGCGTGGAGTTCGACGA 539
QY 268 -----ValGluValLeuGluCysLysIleGlnCysGluGluAsnLe 281
Db 538 CAACACGCTCTGGGCGCACTGACCATCTCGAGCTGAAGTG----- 496
QY 281 uThrProValIleGlyGlyTyrProValGluLysPheValAlaThrMetTyrHisTyrLe 301
Db 495 -----CTCATCGGC----- 487
QY 301 uGlnPheAlaTyrTyrLysLeuAsnAspLeuLysAsnAlaAlaProCysAlaValSerTy 321
Db 486 -----CTCGCCCTGAAGCGCTTCGAGACGCCAAGGAGGCGTGGAGGCGTCTCCAGTA 431
QY 321 rLeuLeuPheAspGlnAsnAspLysValMetGlnGlnAsnLeuValTyrGlnTyrHi 341
Db 430 C-----AACGACACACGCTCGAGCGCGGCTGTCTAC----- 397
QY 341 sArgAspThrTyrGlyLeuSerAspGluHisPheGlnProArgProGluAlaValGlnPh 361
Db 396 -----CAGGC 392
QY 361 ePheAsnValThrThrLeuGlnLysGluLeuTyrAspPheAlaLysGluAsnIleMetAs 381
Db 391 GCTGACAGTGTGCTG-----GAGGTGCTCT 365
QY 381 pAspAspGluGlyGluValValGluTyr 390
Db 364 CGACGACGAGCTGGAGATCGCCGACTAC 337
```

Search completed: January 30, 2004, 08:13:11
Job time : 1055 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 30, 2004, 03:51:33 ; Search time 2712 Seconds
(without alignments)
3593.693 Million cell updates/sec

Title: US-09-729-674-2
Perfect score: 2130
Sequence: 1 NEPGRGAALLALCVACA.....DDEGEVVEYDDILLEETS 401

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 22781392 seqs, 12152238056 residues
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_p2n.model -DEV=xlh
-Q=/cgn2_1/USPFO_spool/US09729674/runat_29012004_095314_12897/app_query.fasta_1.583
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LST=45
-LOCALALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09729674 @CGN 1 1 2810 @runat_29012004_095314_12897 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT_DSBLCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: em estba.*
2: em esthum.*
3: em estin.*
4: em estmu.*
5: em estov.*
6: em estpl.*
7: em estro.*
8: em htc.*
9: gb est1.*
10: gb est2.*
11: gb htc.*
12: gb est3.*
13: gb est4.*
14: gb est5.*
15: em estfun.*
16: em estom.*
17: em gss hum.*
18: em gss inv.*
19: em gss pln.*
20: em gss vrt.*
21: em gss fun.*
22: em gss nam.*
23: em gss mus.*
24: em gss pro.*
25: em gss rod.*
26: em gss phg.*
27: em gss vrl.*
28: gb gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1893	88.9	1656	11	AK047506	AK047506 Mus muscu
2	1868	87.6	1852	11	AK017797	AK017797 Mus muscu
3	1791.5	84.1	1201	9	AL545483	AL545483 AL545483
4	1684	79.1	1201	9	AL556976	AL556976 AL556976
5	1680.5	78.9	1075	9	AL546910	AL546910 AL546910
6	1663	78.1	1148	13	BU902192	BU902192 AGENCOURT
7	1618.5	76.0	1201	13	BU425905	BU425905 BX425905
8	1610	75.6	1201	9	AL517455	AL517455 AL517455
9	1600	75.1	1201	9	AL541167	AL541167 AL541167
10	1584	74.4	1201	9	AL544488	AL544488 AL544488
11	1522.5	71.5	1201	9	AL543090	AL543090 AL543090
12	1498	70.3	1201	9	AL554750	AL554750 AL554750
13	1468	68.9	1430	11	AK007624	AK007624 Mus muscu
14	1426.5	67.0	846	13	BU838140	BU838140 AGENCOURT
15	1419	66.6	1109	12	BM550012	BM550012 AGENCOURT
16	1413.5	66.4	1201	9	AL558186	AL558186 AL558186
17	1406	66.0	1152	9	AL550966	AL550966 AL550966
18	1401.5	65.8	944	13	BU856662	BU856662 AGENCOURT
19	1397	65.6	939	13	BX386576	BX386576 BX386576
20	1394.5	65.5	1083	12	BM920687	BM920687 AGENCOURT
21	1369	64.3	774	9	AL598447	AL598447 DKF2P313G
22	1354	63.6	846	12	B1253401	B1253401 602573711
23	1352	63.5	1040	9	AL558966	AL558966 AL558966
24	1342	63.0	899	13	BQ437797	BQ437797 AGENCOURT
25	1341.5	63.0	962	14	CD516948	CD516948 AGENCOURT
26	1337.5	62.8	904	13	BU931550	BU931550 AGENCOURT
27	1325.5	62.2	1103	13	BQ233085	BQ233085 AGENCOURT
28	1325	62.2	1200	13	BU902214	BU902214 AGENCOURT
29	1298	60.9	878	13	BU183287	BU183287 AGENCOURT
30	1285.5	60.4	945	13	BQ677531	BQ677531 AGENCOURT
31	1285	60.3	827	12	B1760650	B1760650 603044832
32	1277	60.0	875	14	CD359485	CD359485 AGENCOURT
33	1277	60.0	912	13	BU175306	BU175306 AGENCOURT
34	1262	59.2	1201	13	BX437872	BX437872 BX437872
35	1261	59.2	887	13	BQ718961	BQ718961 AGENCOURT
36	1246	58.5	837	13	BQ428356	BQ428356 AGENCOURT
37	1244.5	58.4	782	10	B8877856	B8877856 601489333
38	1241	58.3	713	10	BE746070	BE746070 601578609
39	1233	57.9	846	13	BU189067	BU189067 AGENCOURT
40	1228	57.7	707	10	BG747825	BG747825 602705343
41	1215.5	57.1	1201	13	BX460044	BX460044 BX460044
42	1214	57.0	949	10	BG323659	BG323659 602421942
43	1211	56.9	754	14	CA322247	CA322247 UI-N-FX0-
44	1209	56.8	943	13	BU539494	BU539494 AGENCOURT
45	1198.5	56.3	1088	10	BF025893	BF025893 601670066

ALIGNMENTS

RESULT 1
AK047506
LOCUS
DEFINITION
AK047506 1656 bp mRNA linear HTC 05-DEC-2002
Mus musculus 10 days neonate cerebellum cDNA, RIKEN full-length
enriched library, clone:B930079M13 product:cartilage associated
protein, full insert sequence.
ACCESSION
AK047506
VERSION
AK047506.1 GI:26338809
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Genoscope - Centre National de Sequencage
Bp 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 2233.f For
more information about this cluster, see
<http://www.genoscope.cns.fr/>
<http://cgibin/cluster.cgi?seq=CSDDI02CH09p1&cluster=2233.f>. Contact :
Feng Liang Email : liangliang@life.com URL :
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSDDI02CH09p1.


```

MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
21085660
11217851
The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 1430)
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
Furuno, M., Hanagaki, T., Harai, A., Hayatsu, N., Hiramoto, K.,
Hirooka, T., Hori, E., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,
Suzuki, H., Tagami, M., Tegawa, A., Takahashi, F., Tanaka, T.,
Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K.,
Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
Please visit our web site (http://genome.gsc.riken.go.jp/) for
further details.
CDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5', GAGAGAGAGAGATCAAGAGCTCTTTTCTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. cDNA went
through one round of subtraction to Rot = 20.0. Second strand cDNA
was prepared with the primer adapter of sequence [5',
GAGAGAGAGAGCGCGCAATTAATCTCGAGTTAATAATTAATCTCCCGCCCCC 3']. cDNA
was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI, 3'
end: SstI. Host: SOLR.
FEATURES
Source
1..1430
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:1810027014"
/db_xref="MGI:1893348"
/db_xref="taxon:10090"
/clone="1810027014"
/sex="male"
/tissue_type="pancreas"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 day old"
misc_feature
<1..1037
/note="cartilage associated protein (MGD|MGI:1891221,
GB|NM_019922, evidence: ELASTIN, 99%, match=1655)
putative"
BASE COUNT
322 a 411 c 365 g 332 t
ORIGIN
Alignment Scores:
Pred. No.: 1,248-155 Length: 1430
Score: 1468.00 Matches: 300
Percent Similarity: 91.38% Conservative: 18
Best Local Similarity: 86.21% Mismatches: 27
Query Match: 68.92% Indels: 8
DB: 11 Gaps: 2

```

US-09-729-674-2 (1-401) x AK007624 (1-1430)

Qy 56 SerClyGluHisTrpAlaGluSer***GlyTyrLeuGluLeuSerLeuArgLeuHisArg 75
Db 1 AGCGCCAGGACACTGGCGGAGCGTGGGTACTCTGAGGTAGCTGCGCTGCGCCGCT 60
Qy 76 LeuLeuArgAspSerGluAlaPheCysHisArgAsnGlnSerAlaAla---ProGlnPro 94
Db 61 CTGCTGCGGACAG-AAGGCTTCTGCCACCGCACTGCGAGCGCGCCACCGCAGCACCC 119
Qy 95 GluProAlaAlaGlyLeuAlaSerTyrProGluLeuArgLeuPheGlyGlyLeuArg 114
Db 120 GCGGCC---GCCGCGCGCGGACCGCCAGCACTACGCTCTTCGCGAGCGTGGCGCC 176
Qy 115 ArgAlaHisCysLeuLysArgCysLysGlnGlyLeuProAlaPheArgGlnSerGlnPro 134
Db 177 CGCGCGGTGGCTCAAGCGCTGCAGAG-GGCTGCGCGCTTCGCGAGTCGCGAGCCC 235
Qy 135 SerArgGluValLeuAlaAspPheGlnArgArgGluProTyrLysPheLeuGlnPheAla 154
Db 236 AGCGCTCAGTGTGGCGGACTTTCAGCAGCG-GAGCCCTACCAAGTTTCTGCAAGTTCGCC 294
Qy 155 TyrPheLysAlaAsnAsnLeuProLysAlaAlaAlaAlaAlaHisThrPheLeuLeuLys 174
Db 295 TACTTCAAGCCCAATGACCTCCCGAGGCC-ATCGCTGCGGCTCACACCTATCTCTGAAG 353
Qy 175 HisProAspAspGluMetMetLysArgAsnMetAlaTyrTyrLysSerLeuProGlyAla 194
Db 354 GATCCAGATGACGAGATGATGAAGAAACATGAGGATATTATTAAGAGCTTGCCTGGAGCC 413
Qy 195 GluAspTyrLysAspLeuGluThrLysSerTyrGluSerLeuPheLeuAlaVal 214
Db 414 GAGGACCACTAATAAGACTTGGAAACCAAGCTGTACGAGAGCGCTGTTGTCGCTGGCGCT 473
Qy 215 ArgAlaTyrAsnGlyGluAsnTrpArgThrSerIleThrAspMetGluLeuAlaLeuPro 234
Db 474 CGGCGCTACAAACGGGAGAACTGGAGAACGTCGATTTCGACATCGAGCTCCCGCTCCC 533
Qy 235 AspPheLysAlaPheTyrGluCysLeuAlaAlaCysGluCly-SerArgGluLeuLys 254
Db 534 GACTTCCTCAAGCGCTTCTACGAGTGCCTGCGCTGCGAGGCGCTCCCGGAGATCAA 593
Qy 254 sAspPheLysAspPheTyrLeuSerIleAlaAspHisTyrValGluValLeuGluCysLy 274
Db 594 GGACTTCAA-GACTTCTACCTGTCATAGCAGATCACTATGTGGGAGTCTGGAGTGTA 652
Qy 274 sIleGlnCysGluGluAsnLeuThrProValIleGlyTyrProValGluLysPheVa 294
Db 653 GATTCTGTGTGAGGAGACCCCTCACCCAGTCATAGGAGGCTATCCCGTGAGAAATTTGT 712
Qy 294 lAlaThrMetTyrHisTyrLeuGlnPheAlaTyrTyrLysLeuAsnAspLysAsnAl 314
Db 713 GCGGACCATGTATACACTATTACAGTTTGGTATTACAAAGTTGATGATCTGNAGATGC 772
Qy 314 aAlaProCysAlaValSerTyrLeuLeuPheAspGlnAsnAspLysValMetGlnGlnAs 334
Db 773 AGCCCGTGTGCGCTCAGCTACCTGCTCTTTGACACAGAGTGCAGCGTATGCAACAGAA 832
Qy 334 nLeuValTyrTyrGlnTyrHisArgAspThrTrpGlyLeuSerAspGluHisPheGlnPr 354
Db 833 CTGTGTTACTATCAGTACCACCGGACAAAGTGGGCGCTCTCGGATGAGACCTTCAGCC 892
Qy 354 oArgProGluAlaValGlnPhePheAsnValThrThrLeuGlnLysGluLeuTyrAspPh 374
Db 893 CAGACCCGAGAGAGTTCAGTTCTTTAATGTGACGACGCTCCAGAGGAATGTGACACTT 952
Qy 374 eAlaLysGluAsnIleMetAspAspGluGlyGluValValGluTyrValAspAspLe 394
Db 953 CGCTCAGAAACACCTAATGATGATGAGGAGAGGTTGTGGAGTATGTGGACGACTT 1012
Qy 394 uLeuGluLeuGluGluThrSer 401
Db 1013 GTTGGAGACCGAAGAGTCTGCC 1034

```

RESULT 14
BUB38140
LOCUS AGENCOURT_8071749 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6088309
DEFINITION 5', mRNA sequence.
ACCESSION BUB38140
VERSION BUB38140.1 GI:24022535
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 846)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2326 row: e column: 14
High quality sequence stop: 684.
Location/Qualifiers
1..846
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6088309"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC_112"
/note="Organ: skin; Vector: pOTS7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH MGC Library."
BASE COUNT 199 a 255 c 215 g 177 t
ORIGIN
Alignment Scores:
Pred. No.: 2,85e-151 Length: 846
Score: 1426.50 Matches: 269
Percent Similarity: 96.11% Conservative: 3
Best Local Similarity: 95.05% Mismatches: 2
Query Match: 66.97% Indels: 9
DB: 13 Gaps: 1
US-09-729-674-2 (1-401) x BUB38140 (1-846)
QY 66 TyrLeuGluIleSerLeuArgLeuHisArgLeuLeuArgAspSerGluAlaPheCysHis 85
Db 2 TACCTGGAGNTAGCTGGCTGGCTGACCGCTTCTGGCGCAGCGAGGCTTCTGCCAC 61
QY 86 ArgAsnCysSerAlaAlaProGlnProGluProAlaAlaGlyLeuAlaSerTyrProGlu 105
Db 62 CGCACTGACGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCTCGCCAGTATCCGAG 121
QY 106 LeuArgLeuPheGlyGlyLeuLeuArgArgAlaHisCysLeuLeuArgCysIleGly 125
Db 122 CTGGCGCTCTTCGGGGGCGCTGCTGGCGCGCGCGCGCGCGCGCGCTCGAAGCAGGCG 181
QY 126 LeuProAlaPheArgGlnSerGlnProSerArgGluValLeuAlaAspPheGlnArg 145
Db 182 CTGCCAGCTTCGGCGAGTCCCGAGCCCGAGCGCGCGCGCGCGCGCGCTTCCAGCGCGC 241

```

146 GluProTyrLysPheLeuGlnPheAlaTyrPheLysAlaAsnAsnLeuProLysAlaIle 165
242 GAGCCCTCAAGTCTCTGAGTTCGTTACTTCAAGGCAATAATCTCCCAAGGCATC 301
QY 166 AlaAlaAlaHisThrPheLeuLeuLysHisProAspAspGluMetMetLysArgAsnMet 185
Db 302 GCGCTGCTCAACACCTTTCTACTGAAGCATCTCTGATGAGCAATGATGAAGAGGACATG 361
QY 186 AlaTyrTyrLysSerLeuProGlyAlaGluAspTyrIleLysAspLeuGluThrLysSer 205
Db 362 GCATATTATAAGACCTGCTGTCGCGAGGACTACATTAAAGACCTGGAACCAAGTCA 421
QY 206 TyrGluSerLeuPheIleArgAlaValArgAlaTyrAsnGlyGluAsnTrpArgThrSer 225
Db 422 TATGAAGGCTGTTCATCCGAGCGGCGGATCAACGGTGAGAACCTGGAGAACATCC 481
QY 226 IleThrAspMetGluLeuAlaLeuProAspPhePheLysAlaPheTyrGluCysLeuAla 245
Db 482 ATCAGACATCGAGCTGCGCTTCCGACTTCTTCAAGCCTTTACGAGTGTCTCGCA 541
QY 246 AlaCysGluGlySerArgGluLeuLysAspPheLysAspPheTyrLysSerIleAlaAsp 265
Db 542 GCCTGGCGGGTTCCAGGAGATCAAGGACTTCAAGGATTTCTACCTTTCCATAGCAGAT 601
QY 266 HisTyrValGluValLeuGluCysLysIleGlnCysGluGluAsnLeuThrProValIle 285
Db 602 CATTATGTAGAGTTCTGGAATGCAAAATACAGTGTCAAGAGAACCTCACCCAGTTATA 661
QY 286 GlyGlyTyrProValGluLysPheValAlaThrMetTyrHisTyrLeuGlnPheAlaTyr 305
Db 662 GGAGGCTATCCGGTTGAGAAATTTGGCTTACCATGTATCATTTACGAGTTTGCCTAT 721
QY 306 TyrLysLeuAsnAspLeuLysAsnAlaAlaProCysAlaValSerTyrLeuLeuPheAsp 325
Db 722 TATAAGTTGAACACCTGAAGATGAGGCCCTGTGSCAGTCAGCTATCTGCTCTTTGAT 781
QY 326 GlnAsnAspLysValMetGlnGlnAsnLeuValTyrTyrGlnTyrHisArgAspThrTrp 345
Db 782 CAGATGACAAAGTCTATGCAA-----GCAGAACCTTGG 814
QY 346 GlyLeuSer 348
Db 815 GGTATTACC 823

RESULT 15
BUB50012
LOCUS AGENCOURT_6559241 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:5443288
DEFINITION 5', mRNA sequence.
ACCESSION BUB50012
VERSION BUB50012.1 GI:18785873
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1109)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1920 row: i column: 17
High quality sequence start: 20
High quality sequence stop: 714.

FEATURES

Location/Qualifiers

1. .1109
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5443288"
 /tissue_type="astrocytoma grade IV, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_98"
 /note="Organ: Brain; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 277 a 331 c 282 g 219 t

ORIGIN

Alignment Scores:

Pred. No.: 3.1e-150 Length: 1109
 Score: 1419.00 Matches: 269
 Percent Similarity: 98.55% Conservative: 3
 Best Local Similarity: 97.46% Mismatches: 1
 Query Match: 66.62% Indels: 3
 DB: 12 Gaps: 0

US-09-729-674-2 (1-401) x BM550012 (1-1109)

QY 84 CysHisArgAsnCyssSerAlaAlaProGlnProGluProAlaAlaGlyLeuAlaSerTyr 103
 Db 33 TGCCACCGCACTGTAGCCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCTAT 92
 QY 104 ProGluLeuArgLeuPheGlyGlyLeuLeuArgArgAlaHisCysLeuLysArgCysLys 123
 Db 93 CCGAGCTGCGCTCTTCGGGGGCGCTGCTCGCGCGCGCGCGCGCGCGCGCGCGCAAG 152
 QY 124 GlnGlyLeuProAlaPheArgGlnSerGlnProSerArgGluValLeuAlaAspPheGln 143
 Db 153 CAGGGCTGCCAGCTTCGCCAGTCCAGCCCGCGCGCGCGCGCGCGCGCGCGCGCTCCAG 212
 QY 144 ArgArgGluProTyrLysPheLeuGlnPheAlaTyrPheLysAlaAsnLeuProLys 163
 Db 213 CGCCGGGAGCCCTACAAAGTTCTTCGAGTTCGTTACTTCAAGCAAAATAATCTCCCAA 272
 QY 164 AlaAlaAlaAlaHisThrPheLeuLeuLysHisProAspAspGluMetMetLysArg 183
 Db 273 GCCATCGCCGCTCTCACACCTTCTACTGAGCATCCTGTATGAGCAATGATGAGAGG 332
 QY 184 AsnMetAlaTyrTyrLysSerLeuProGlyAlaGluAspTyrIleLysAspLeuGluThr 203
 Db 333 AACATGGCATATTATAAGAGCTGCTGCTGCGGAGGACTACATTAAAGACCTGGAAACC 392
 QY 204 LysSerTyrGluSerLeuPheLeuArgAlaValArgAlaTyrAsnGlyGluAsnTrpArg 223
 Db 393 AAGTCATATGAAGCCTGTTCATCCGAGCAGTCCGGGCGCATCAACGGTGAGAACTGGAGA 452
 QY 224 ThrSerIleThrAspMetGluLeuAlaLeuProAspPhePheLysAlaPheTyrGluCys 243
 Db 453 ACATCCATCACAGACATGGAGCTGGCCCTTCCGACTTCTTCAAGCCCTTTACGAGTGT 512
 QY 244 LeuAlaAlaCysGluGlySerArgGluIleLysAspPheLysAspPheTyrLeuSerIle 263
 Db 513 CTCGCAGCCTGCCAGGGTTCCAGGGAGATCAAGGACTTCAAGGATTTCTACCTTTCCATA 572
 QY 264 AlaAspHisTyrValGluValLeuGluCysLysIleGlnCysGluGluAsnLeuThrPro 283
 Db 573 GCGAGATCATTTAGTAGAGTTCTGGAAATGCAAAATACAGTGTGAGAGAACCTCACCCA 632
 QY 284 ValIleGlyGlyTyrProValGluLysPheValAlaThrMetTyrHisTyrLeuGlnPhe 303
 Db 633 GTTATAGGAGGCTATCCGGTTGAGAAATTTGTGGCTACCATTGATTCATTGCGAGTTT 692

QY 304 AlaTyrTyrLysLeuAsnAspLeuLysAsnAlaAlaProCysAlaValSerTyrLeuLeu 323
 Db 693 GCCTATTATTAAGTTGAACGACCTGAAGAATGCAGCCCCCTGTGCAGTCAGCTATCTGCTC 752
 QY 324 PheAspGlnAsnAspLysValMetGlnGlnAsnLeuValTyrTyrGlnTyrHisArgAsp 343
 Db 753 TTTGATCAGAAATGAACAGGTTCATGCAGCAGAACCTGTTGTTATTACCAGTACCACAGGGAA 812
 QY 344 -ThrTrp-GlyLeuSerAspGluHisPhe-GlnProArgPro 356
 Db 813 CACTTGGGGGCGCTCTCGGATGAGCACCTCCAGCCCCAGACCC 854

Search completed: January 30, 2004, 07:55:45
 Job time : 2729 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 30, 2004, 03:39:47 ; Search time 357 Seconds
(without alignments)
3032.142 Million cell updates/sec

Title: US-09-729-674-2

Perfect score: 2130

Sequence: 1 MFPGRRGAALALLICVACA.....DDEGEWVEYVDLLELETS 401

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5

Ygapop 10.0, Ygapext 0.5

Fgapop 6.0, Fgapext 7.0

Delop 6.0, Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-Q=/cpn2_1/USPTO.spool/US09729674/runat_29012004_095313_12879/app_query.fasta_1.583
-DB=N_Geneseq_19Jun03 -QFMT=fastep -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORW=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09729674 @CN1_1.0 @runat_29012004_095313_12879 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -D5PBLOCK=100 -LONGLOG
-DSV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -FGAPEXT=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_19Jun03.*

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	2128	99.9	3871	20	Human secreted pro
2	2128	99.9	3871	22	Human cDNA encodin
3	2128	99.9	3871	24	Human polynucleoti
4	2121	99.6	1961	22	Human cDNA sequenc
5	1247	58.5	1586	21	Human colon cancer
6	1155	54.2	818	22	Human cDNA clone (
7	1088.5	51.1	2347	24	Cesophagus cancer
8	717	33.7	408	21	Human colon cancer
9	574	26.9	2242	24	Human polynucleoti
10	572.5	26.9	2993	22	Human cDNA sequenc
11	571	26.8	2211	24	DNA encoding novel
12	571	26.8	2524	22	Human cDNA sequenc
13	571	26.8	2563	22	Human cDNA encodin
14	571	26.8	2585	22	Human cDNA sequenc
15	568	26.7	1309	24	Human polynucleoti
16	567	26.6	2600	21	Human tumour suppr
17	562.5	26.4	2577	22	Human secreted pro
18	549	25.8	2152	24	Human polynucleoti
19	543.5	25.5	2615	22	Human secreted pro
20	539	25.3	2829	21	Human tumour suppr
21	538	25.3	2322	21	Mouse tumour suppr
22	538	25.3	2416	21	Human novel polynu
23	527	24.7	2753	24	Human novel polynu
24	527	24.7	2753	24	Human novel polynu
25	473.5	22.2	2127	24	DNA encoding novel
26	456	21.4	2665	25	cDNA downregulated
27	382.5	18.0	2789	22	Human cDNA sequenc
28	376.5	17.7	800	22	Human cDNA clone (
29	340.5	15.5	1025	24	Oligonucleotide fo
30	330.5	15.5	1025	24	Oligonucleotide fo
31	321.5	15.1	2305	25	DNA encoding human
32	306.5	14.4	1026	24	Oligonucleotide fo
33	306.5	14.4	1026	24	Oligonucleotide fo
34	302	14.2	550	25	Human adipocyte se
35	283.5	13.3	774	22	Human cDNA clone (
36	273.5	12.8	652	22	Primer specific fo
37	273.5	12.8	703	22	Human cDNA clone (
38	268	12.6	1026	24	Oligonucleotide fo
39	268	12.6	1026	24	Oligonucleotide fo
40	267	12.5	1025	24	Oligonucleotide fo
41	267	12.5	1025	24	Oligonucleotide fo
42	285.5	12.5	738	22	Human cDNA clone (
43	252	11.8	963	23	DNA encoding novel
44	214.5	10.1	682	22	Human neuroblastom
45	200.5	9.4	349	21	Lung cancer associ

ALIGNMENTS

RESULT 1

AAK60801

ID AAK60801 standard; DNA; 3871 BP.

XX AAK60801;

XX 09-AUG-1999 (first entry)

DT Human secreted protein encoding DNA (clone bd306-7).

DE Secreted protein; kidney; lung; brain; blood; testis; bone marrow;
KW nutritional activity; cytokine; cell proliferation; immune stimulation;
KW hematopoiesis regulation; tissue growth; thrombolytic; gene therapy;
KW anti-inflammatory; tumour invasion; ss.

XX Homo sapiens.

OS XX


```
QY 121 ArgCysLysGlnGlyLeuProAlaPheArgGlnSerGlnProSerArgGluValLeuAla 140
Db 423 CGCTGCAAGCAGGCGCTGCAGCCTTCGCCAGTCCAGCCAGCGCGAGGTGTGCGG 482
QY 141 AspPheGlnArgArgGluProTyrLysPheLeuGlnPheAlaTyrPheLysAlaAsn 160
Db 483 GACTTCCAGCGCGCGAGCCCTACAGTTCCTGCAGTTCGCTTACTTCAAGCAATAAT 542
QY 161 LeuProLysAlaAlaAlaAlaAlaHisThrPheLeuLeuLysHisProAspAspGluMet 180
Db 543 CTCGCCAAAGCCATCCCGCTGCTCACACCTTCTACTGAAGCATCCTGATGACGAAATG 602
QY 181 MetLysArgAsnMetAlaTyrTyrLysSerLeuProGlyValAlaGluAspTyrLysAsp 200
Db 603 ATGAAGAGGAACATGSCATATATATAGAGCCTGCCTGTCGCGAGACTACATTAAAGAC 662
QY 201 LeuGluThrLysSerTyrGluSerLeuPheLeuAlaValArgAlaTyrAsnGlyGlu 220
Db 663 CTGGAACCAAGTCATATGAAGCCTGTTTCATCCGAGCAGTCCGCGCATACAAACGTCAG 722
QY 221 AsnTrpArgThrSerIleThrAspMetGluLeuAlaLeuProAspPhePheLysAlaPhe 240
Db 723 AACTGGAGAACATCCATCATCAGACATGAGCTGGCCCTTCCCGATCTCTTCAAGCCCTT 782
QY 241 TyrGluCysLeuAlaAlaCysGluGlySerArgGluIleLysAspPheLysAspPheTyr 260
Db 783 TAGAGTGTCTCGAGCCTGCAGGGTTCAGGGAGATCAAGGACTTCAAGGATTTCTAC 842
QY 261 LeuSerIleAlaAspHisTyrValGluValLeuGluCysLysIleGlnCysGluGluAsn 280
Db 843 CTTTCATAGCAGATCATATATGAAGTCTGGAATGCAAAATACAGTGTGAAGAGAAC 902
QY 281 LeuThrProValIleGlyTyrProValGluLysPheValAlaThrMetTyrHisTyr 300
Db 903 CTCACCCAGTATAGGAGCTATCCGTTGAGAAATTTGGCTACCATGTATCATTTAC 962
QY 301 LeuGlnPheAlaTyrTyrLysLeuAsnAspLeuLysAsnAlaAlaProCysAlaValSer 320
Db 963 TTGCACTTTGCCCTATTATTAAGTTGAACGACCTGAAGAATGCGAGCCCTGTGCAGTCAGC 1022
QY 321 TyrLeuLeuPheAspGlnAsnAspLysValMetGlnGlnAsnLeuValTyrGlnTyr 340
Db 1023 TATCTGCTTTTATCAGAAATGACAGGTCTATGAGAGAACTCGTGTATTACCAAGTAC 1082
QY 341 HisArgAspThrTrpGlyLeuSerAspGluHisPheGlnProArgProGluAlaValGln 360
Db 1083 CACAGGACACTTGGGCGCTCTCGATGAGCATTCCAGCCAGACCTGAAGCAGCTTCAG 1142
QY 361 PhePheAsnValThrThrLeuGlnLysGluLeuTyrAspPheAlaLysGluAsnIleMet 380
Db 1143 TTCTTTAATGTACCACACTCCAGAGGAGCTGTATGACTTTGCTAAGGAAATATAATG 1202
QY 381 AspAspAspGluGlyGluValValGluTyrValAspAspLeuLeuGluLeuGluThr 400
Db 1203 GATGATGATGAGGAGAACTTGTGGAATATGTGGATGACCTTGTGAACTTGGAGGAGACC 1262
QY 401 Ser 401
Db 1263 AGC 1265
RESULT 4
AAH14525
ID AAH14525 standard; cDNA; 1961 BP.
XX AAH14525;
AC
XX
XX
XX 26-JUN-2001 (first entry)
XX Human cDNA sequence SEQ ID NO:12066.
DE Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
KW
XX
```

```
OS Homo sapiens.
XX EP1074617-A2.
XX 07-FEB-2001.
XX 28-JUL-2000; 2000EP-0116126.
XX 29-JUL-1999; 99JP-0248036.
XX 27-AUG-1999; 99JP-0300053.
XX 11-JAN-2000; 2000JP-0118776.
XX 02-MAY-2000; 2000JP-0183767.
XX 09-JUN-2000; 2000JP-0241899.
XX (HELI-) HELIX RES INST.
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX Primer sets for synthesizing polynucleotides, particularly the 5602
XX full-length cDNAs defined in the specification, and for the detection
XX and/or diagnosis of the abnormality of the proteins encoded by the
XX full-length cDNAs -
XX Claim 8; SEQ ID 12066; 2537pp + CD ROM; English.
XX The present invention describes primer sets for synthesizing 5602
XX full-length cDNAs defined in the specification. Where a primer set
XX comprises: (a) an oligo-dr primer and an oligonucleotide complementary
XX to the complementary strand of a polynucleotide which comprises one of
XX the 5602 nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises a 3'-end sequence, where the
XX oligonucleotide comprises at least 15 nucleotides and the combination
XX of the 5'-end sequence/3'-end sequence is selected from those defined in
XX the specification. The primer sets can be used in antisense therapy and
XX particularly full-length cDNAs. The primers are also useful for the
XX detection and/or diagnosis of the abnormality of the proteins encoded by
XX the full-length cDNAs. The primers allow obtaining of the full-length
XX cDNAs easily without any specialized methods. AAH03166 to AAH13628 and
XX AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
XX AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
XX represent oligonucleotides, all of which are used in the exemplification
XX of the present invention.
SQ Sequence 1961 BP; 417 A; 579 C; 481 G; 484 T; 0 other;
Alignment Scores:
Pred No.: 1-35e-200 Length: 1961
Score: 2121.00 Matches: 399
Percent Similarity: 99.50% Conservative: 0
Best Local Similarity: 99.50% Mismatches: 2
Query Match: 99.58% Indels: 0
DB: 22 Gaps: 0
US-09-729-674-2 (1-401) x AAH14525 (1-1961)
QY 1 MetGluProGlyArgArgGlyAlaAlaAlaLeuLeuAlaLeuLeuCysValAlaCysAla 20
Db 40 ATGAGCGCGGCGCGCGCGCGCGCGCGCGCGCTGCTAGCGCTGCTGCGCTGCGCTGCGG 99
QY 21 LeuArgAlaGlyArgAlaGlnTyrGluArgTyrSerPheArgSerPheProArgAspGlu 40
Db 100 CTGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCTAGCGCTGCTGCGCTGCGCTGCGG 159
QY 41 LeuMetProLeuGluSerAlaTyrArgHisAlaLeuAspLysTyrSerGlyGluHisTrp 60
Db 100 CTGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCTAGCGCTGCTGCGCTGCGCTGCGG 159
```

Db 160 CTGATGCCGCTCGAGTCCGCTACCGGCACGCGCTGGGCAAGTACAGCGCGGAGCACTGG 219

Qy 61 AlaGluSer**GlyTyrLeuGluLeuHisArgLeuLeuArgAspSer 80

Db 220 GCGAGAGCGTGGGCTACCTGGAGATCAGCTTCGGGCTGCACCGCTTGTCTGGCGCAGC 279

Qy 81 GluAlaPheCysHisArgAsnCysSerAlaAlaProGlnProGluProAlaAlaGlyLeu 100

Db 280 GAGGCCCTTCGCCACCGCACTGAGCGCGCGCGCGCAGCGCGCGCGCGCGCGCTC 339

Qy 101 AlaSerTyrProGluLeuArgLeuPheGlyGlyLeuLeuArgArgAlaHisCysLeuLys 120

Db 340 GCCAGCTATCCGAGCTCGCCCTTCGCGGGGCTCTGCGCGCGCGCACTGCCTCAG 399

Qy 121 ArgCysLysGlnGlyLeuProAlaPheArgGlnSerGlnProSerArgGluValLeuAla 140

Db 400 CGTGCAGCAGGCGCTGCCAGCGCTTCGCGCAGTCCCGAGCGCGCGCGCGCGCTGCGG 459

Qy 141 AspPheGlnArgArgGluProTyrLysPheLeuGlnPheAlaTyrPheLysAlaAsnAsn 160

Db 460 GACTTCCAGCGCGCGAGCGCTACAGTTCTCGAGTTCTGCTTACTTCAAGGCAATAT 519

Qy 161 LeuProLysAlaLeuAlaAlaHisThrPheLeuLeuLysHisProAspAspGluMet 180

Db 520 CTCGCCAAAGCATCGCGCTGCTGCACCTTCTACTGAAGCATCTGATGACGAAATG 579

Qy 181 MetLysArgAsnMetAlaTyrTyrLysSerLeuProGlyAlaGluAspTyrLysAsp 200

Db 580 ATGAAGAGGACATGGCATATTAAGAGCGCTGCTGCGCGCGCGAGCTACATTAAGAC 639

Qy 201 LeuGluThrLysSerTyrGluSerLeuPheLeuAlaValArgAlaTyrAsnGlyGlu 220

Db 640 CTGGAACCAAGTCATATGAAGCGCTTTCATCCGAGCAGTCCGGGCATACACAGGTGAG 699

Qy 221 AsnTrpArgThrSerIleThrAspMetGluLeuAlaLeuProAspPhePheLysAlaPhe 240

Db 700 AACTGGAGACATCCATCAGACATGGAGCTGGCGCTTCCCGACTTCTTCAAGCGCTT 759

Qy 241 TyrGluCysLeuAlaAlaCysGluGlySerArgGluLysAspPheLysAspPheTyr 260

Db 760 TACGAGTGTCTCGCAGCGCTCGAGGTTCCAGGGAGATCAAGGACTTCAAGGATTTCTAC 819

Qy 261 LeuSerIleAlaAspHisTyrValGluValLeuGluCysLysIleGlnCysGluGluAsn 280

Db 820 CTTTCCATAGCAGATCATATTATGAGAAGTTCTGGAATGCAAAATACAGTGTGAAGAGAC 879

Qy 281 LeuThrProValIleGlyGlyTyrProValGluLysPheValAlaThrMetTyrHisTyr 300

Db 880 CTCACCCAGTTATAGAGGCTATCCGGTTGAGAAATTTGGCTACCATGATCATTTAC 939

Qy 301 LeuGlnPheAlaTyrTyrLysLeuAsnAspLeuLysAsnAlaAlaProCysAlaValSer 320

Db 940 TTGCAGTTTGCCTATTATTAAGTTGAACGACTGAAGAATCGAGCCCTGTGCAGTCAGC 999

Qy 321 TyrLeuLeuPheAspGlnAsnAspLysValMetGlnGlnAsnLeuValTyrGlnTyr 340

Db 1000 TATCTGCTTTTGTATCAGATGACAGGTATGCGAGCACTCTGGTGATTAACAGTAC 1059

Qy 341 HisArgAspThrTrpGlyLeuSerAspGluHisPheGlnProArgProGluAlaValGln 360

Db 1060 CACAGGACACGCTGGGCGCTCTCAGATGAGCACTTCCAGCCCGAGCTGAGCAGTTTCA 1119

Qy 361 PhePheAsnValThrThrLeuGlnLysGluLeuTyrAspPheAlaLysGluAsnIleMet 380

Db 1120 TTCTTTTATGTGACCACTCCAGAGGAGCTGTATGACTTGTCTAAGGAAATATAATG 1179

Qy 381 AspAspAspGluGlyGluValValGluTyrValAspAspLeuLeuGluLeuGluThr 400

Db 1180 GATCATGATGAGGAGAGAGTTGTGATATATGTGATGACCTCTTGGNACTGGAGGAGC 1239

Qy 401 Ser 401

Db 1240 AGC 1242

RESULT 5

AAH31138
ID AAH31138 standard; cDNA; 1586 BP.
XX
XX AAH31138;
AC
XX 27-JUL-2001 (first entry)
DT
DE Human colon cancer cell line Km12L4-A cDNA library derived seq#1072.
XX
XX Human; diagnosis; colon cancer; cancer; malignant; chromosome mapping;
KW detection; colon cancer cell line Km12L4-A; ss.
KW
XX Homo sapiens.
OS
XX WC200018916-A2.
PN
XX 06-APR-2000.
PD
XX 23-SEP-1999; 99WO-US22226.
PF
XX 28-SEP-1998; 98US-0102161.
PR 28-SEP-1998; 98US-0102180.
PR 29-SEP-1998; 98US-0102380.
PR 08-OCT-1998; 98US-0103815.
PR 27-OCT-1998; 98US-0105877.
XX
XX (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
PI Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;
PI Lamson G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;
PI Leskowitz D, Kita D, Garcia V, Jones LW, Stache-Crain B;
XX
XX WPI; 2000-299155/25.
DR
XX Polynucleotide library comprising 1079 defined sequences, useful in
PT the form of an array to detect cancer or susceptibility to cancer -
PT
XX Claim 1; Page 498-499; 502pp; English.
PS
XX The present invention describes a library of polynucleotides comprising
CC 1079 nucleotide sequences (given in AAH30067 to AAH31145). Also described
CC are: (1) an isolated polynucleotide (I) having at least 90% identity to
CC one of the 1079 sequences; (2) a recombinant host cell containing (I);
CC (3) an isolated polypeptide (II) encoded by (I); (4) an antibody that
CC specifically binds to (II); (5) a vector comprising (I); and (6) a method
CC of detecting differentially expressed genes correlated with a cancerous
CC state of a mammalian cell comprising detecting a gene product encoded by
CC 65 of the 1079 sequences given in the specification. The polynucleotides
CC are used to monitor patients having (or susceptible) to cancer to detect
CC potentially malignant events at a molecular level before they are
CC detectable at a gross morphological level. The polynucleotides are also
CC useful for monitoring the efficacy of various therapies and preventive
CC interventions. Polynucleotide probes based on the disclosed sequences
CC are useful for chromosome mapping and detection of transcription levels.
CC The 1079 polynucleotide sequences were derived from a human colon cancer
CC cell line Km12L4-A cDNA library.
XX
XX Sequence 1586 BP; 381 A; 406 C; 340 G; 453 T; 6 other;
SQ

Alignment Scores:

Pred. NO.: 6.39e-114 Length: 1586
Score: 1247.00 Matches: 233
Percent Similarity: 99.15% Conservative: 0
Best Local Similarity: 99.15% Mismatches: 2
Query Match: 58.54% Indels: 0
DB: 21 Gaps: 0

US-09-729-674-2 (1-401) X AAH31138 (1-1586)

```
QY 167 AlalaHisThrPheLeuLeuLysHisProAspAspGluMetMetLeuArgAsnMetAla 186
DB 3 GCTGTCTCACACCTTTCTACTAGGATCTCTGATGACGAAATGATGAAGAGAAATGCA 62
QY 187 TyrTyrLysSerLeuProGlyAlaGluAspTyrIleLysAspLeuGluThrLysSerTyr 206
DB 63 TATTATAAGAGCGCTGCTGCTGCCGAGGACTCATTAAGAGCTTGAACCAAGTCATAT 122
QY 207 GluSerLeuPheIleArgAlaValArgAlaTyrAsnGlyGluAsnTrpArgThrSerIle 226
DB 123 GAAAGCCTTTCATCCGAGCAGTGGGCATACAAAGGTGAGAACTGGAGAACATCCATC 182
QY 227 ThrAspMetGluLeuAlaLeuProAspPhePheLysAlaPheTyrGluCysLeuAla 246
DB 183 ACAGACATGGAGCTGCCCTTCCGACTTCTTCAAGGCTTTTACGAGTCTGCGAGCC 242
QY 247 CysGluGlySerArgGluLeuLysAspPheLysAspPheTyrLeuSerIleAlaAspHis 266
DB 243 TCCRRGGGTTCCAGGAGATCAAGGACITCAAGGATTTCTACCTTTCATAGCAGATCAT 302
QY 267 TyrValGluValLeuGluCysLysIleGlnCysGluGluAsnLeuThrProValIleGly 286
DB 303 TATGTAGAGTTCGTGAATGCAAAATACAGTGTGAAGAGAACTCACCCAGTTATAGGA 362
QY 287 GlyTyrProValGluLysPheValAlaThrMetTyrHisTyrLeuGlnPheAlaTyrTyr 306
DB 363 GACTATCCGGTGAATAATTTGGTCTACCATGATATCAATTCTTCAGTTTGCCTATTAT 422
QY 307 LysLeuAsnAspLeuLysAsnAlaProCysAlaValSerTyrLeuLeuPheAspGln 326
DB 423 AAGTTGAACGACCTGAAGAAATGACGCCCTGTCAGTCAGTACTCTCTTTCATCAG 482
QY 327 AsnAspLysValMetGlnGlnAsnLeuValTyrTyrGlnTyrHisArgAspThrTrpGly 346
DB 483 AATGCAAGGTCATGACAGAACCTGGTGATTTACCAAGTACCAAGGAGACACTGGGGC 542
QY 347 LeuSerAspGluHisPheGlnProArgProGluAlaValGlnPheAsnValThrThr 366
DB 543 CTCTCKGATGAGCACTTCAGCCAGCCTGAGCAGTTCAGTTCTTTAATGTGACCACA 602
QY 367 LeuGlnLysGluLeuTyrAspPheAlaLysGluAsnIleMetAspAspGluGlyGlu 386
DB 603 CTTCCAGAGGAGCTGTATGACTTTCTAAGGAAATATAAATGATGATGATGAGGAGAA 662
QY 387 ValValGluTyrValAspAspLeuLeuGluLeuGluGluThrSer 401
DB 663 GTTGTGGATATGTGGATGACCTCTTGGAACTTGGAGGAGCAGC 707

RESULT 6
AAH06808
ID AAH06808 standard; cDNA; 818 BP.
XX
AC AAH06808;
XX
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA clone (5'-primer) SEQ ID NO:3643.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
```

```
XX (HELI-) HELIX RES INST.
PA
XX Ota T, Isozai T, Nishikawa T, Hayaehi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
XX Claim 1; SEQ ID 3643; 2537bp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 818 BP; 159 A; 271 C; 236 G; 148 T; 4 other;

Alignment Scores:
Pred. No.: 3.5e-105 Length: 818
Score: 1155.00 Matches: 238
Percent Similarity: 91.57% Conservative: 1
Best Local Similarity: 91.19% Mismatches: 15
Query Match: 54.23% Indels: 8
DB: 22 Gaps: 1

US-09-729-674-2 (1-401) x AAH06808 (1-818)
QY 1 MetGluProGlyArgArgGlyAlaAlaAlaLeuLeuAlaLeuLeuLeuAlaCysAla 20
DB 40 ATGGAGCGGGGCGCGGGGGCGCGGGCTGCTAGCGCTGCTGCTGCGCTGCGCG 99
QY 21 LeuArgAlaGlyArgAlaGlnTyrGluA:GlyTyrSer:PheArgSerPheProArgAspGlu 40
DB 100 CTGCGCGCGGGGCGCGCCCAATACGAACGCTACAGCTTCGCGAGCTTCCGCGGACGAG 159
QY 41 LeuMetProLeuGluSerAlaTyrArgHisAlaLeuAspLysTyrSerGlyGluHisTyr 60
DB 160 CTGATGCCGCTCGAGTCGGCTTACCGGCGCGCGCTGGCGAGTACAGCGCGGAGCACTGG 219
QY 61 AlaGluSer***GlyTyrLeuGluIleSerLeuArgLeuHisArgLeuLeuArgAspSer 80
DB 220 GCCGAGAGCGTGGGCTACCTGGAGATCAGCTCGCGCTGACCGCTTGTGCGCAGAG 279
QY 81 GluAlaPheCysHisArgAsnCysSerAlaAlaProGlnProGluProAlaAlaGlyLeu 100
DB 280 GAGGCTTCTGCCACCGCAACTGCAGCGCGCGCGCGCGAGCCCGCGCGCGCGCTC 339
QY 101 AlaSerTyrProGluLeuArgLeuPheGlyGlyLeuLeuArgAlaHisCysLeuLys 120
DB 340 GCCAGCTATCCGAGCTGCGCTTTCGGGGGCGCTGCTGCGGGGCGCTGCTGCGCGCAG 399
```


QY 44 LeuGluSerAlaTyrArgHisAlaLeuAspLysTyrSerGlyClnHisTrpAlaGluSer 63
Db 117 CTGGCGCGGGCTACGGGCGCGCTCTGGAGCAGTACGAGGAGAGAGCTGGCGCGAGC 176
QY 64 ***GlyTyrLeuGluLeuSerLeuArgLeuHisArgLeuLeuArgAspSerGluAlaPhe 83
Db 177 GCGCGCTACCTGAGGCGGCTCGGCTGCACCGGCTCTGGCGCAGCAGGCGCTTC 236
QY 84 CysHisArgAsnCysSerAla-----AlaProGlnProGluProAlaAlaGlyLeu 100
Db 237 TGCACGCCCACTGCAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 296
QY 101 AlaSer-----TyrProGluLeuArgLeuPheGlyGlyLeuArgAlaHisCys 118
Db 297 GCAGCAGGCGGCTCGGCTCGGCTCTTCGCGCGGCTCTGGAGCAGCGCGCTTC 356
QY 119 LeuLysArgCysLysGlnGlyLeuProAlaPheArgGlnSerGlnProSerArgGluVal 138
Db 357 CTGGCGGCTGCAAGCGGAGCGCTCGCGCGCTTCGCGCGGCTCTGGAGCAGCGCGCT 416
QY 139 LeuAlaAspPheGlnArgGluProTyrLysPheLeuGlnPheAlaTyrPheLysAla 158
Db 417 CTGCGTGACTTCAGAGCGCGCTCGCGCTTACAGTACCTGCTACCTGCTTCAAGGCT 476
QY 159 AsnAsnLeuProLysAlaLeuAlaAlaHisThrPheLeuLeuLysHisProAspAsp 178
Db 477 AACCGGCTGAGAAGCGGCTGGCGGCGCTACACCTTCTTCAGAGGAAACCGAAGCAC 536
QY 179 GluMetMetLysArgAsnMetAlaTyrTyrLysSer---LeuProGlyAlaGluAspTyr 197
Db 537 GAGCTACCGCGCAAGTATCTCAACTACTATCAGGAGGAGTCTGGAGCTCGCGCGAGTCC 596
QY 198 IleLysAspLeuGluThrLysSerTyrGluSerLeuPheLeuArgAlaValArgAlaTyr 217
Db 597 CTCACGAGCTAGAGCGCGCGCTAGAGCGCGCTTCTCGCGGCTGTGAGCTCTAC 656
QY 218 AsnGlyGluAsnTrpArgThrSerIleThrAspMetGluLeuAlaLeuProAspPhePhe 237
Db 657 AACAGCGGGGATTCGCGCAGCAGCAGCGAGGACATCGAGCGGCGCTTGTACAGTACCTG 716
QY 238 LysAlaPheTyrGluCysLeuAlaAlaCysGluGlySerArgGluLeuLysAspPheLys 257
Db 717 GCAGTCTTTCGCGGCTGCTGCGCGCTGTGAAGGGCGCCATGACAGGTGAGCTTCAAG 776
QY 258 AspPheTyrLeuSerIleAlaAspHisTyrValGluValLeuGluCysLysIleGlnCys 277
Db 777 GACTTCTACCGGCCCATAGCAGATCTTTCGAGAGTCCCTGCAGTGCAAGGTGACGTGT 836
QY 278 GluGluAsnLeuThrProValIleGlyTyrProValGluLysPheValAlaThrMet 297
Db 837 GAGGCCAATTTGACCCCAATGGGTGGCTACTTCTGTGGAAGAAGTCTGTGGCCACCATG 896
QY 298 TyrHisTyrLeuGlnPheAlaTyrTyrLysLeuAsnAspLeuLysAsnAlaAlaProCys 317
Db 897 TACCACCTACCTGAGTTTCCCTACTATAGTTGATGTGCGCGCGAGCTGCGCGCAGC 956
QY 318 AlaValSerTyrLeuLeuPheAspGlnAsnAspLysValMetGlnGlnAsnLeuValTyr 337
Db 957 GCGCGCAGCTACATGTCTTCGACCCCAAGCAGCAGCGCTCATGCAAGAACCTGTGTAT 1016
QY 338 TyrGlnTyrHisArgAspThrTrpGlyLeuSerAspGluHisPheGlnProArgProGlu 357
Db 1017 TACCGTTTCCACCGGCTGCTGGGCGCTGGAGAGGAGGAGTTCAGCCCGCGGAGAG 1076
QY 358 AlaValGlnPhePheAsnValThrThrLeuGlnLysGluLeuTyrAspPheAlaLysGlu 377
Db 1077 GCATGCTCTACCAACACAGACCGCGCGAGCTGCGGAGCTGCTGGAGTTCACCCACATG 1136
QY 378 AsnIleMetAspAspAspGluGlyGluValValGluTyrValAspAspLeuLeuGluLeu 397
Db 1137 TACCTGCATCATATGATGAG-----ATGAGGCTG 1166
QY 398 GluGluThr 400

Db 1167 GAGGAGACA 1175

RESULT 8

AAH30585
ID AAH30585 standard; cDNA; 408 BP.

AC AAH30585;

DT 27-JUL-2001 (first entry)

DE Human colon cancer cell line Km12L4-A cDNA library derived sequence #519.

DE Human; diagnosis; colon cancer; cancer; malignant; chromosome mapping;
KW detection; colon cancer cell line Km12L4-A; ss.

XX Homo sapiens.

XX WO200018916-A2.

XX 06-APR-2000.

XX 23-SEP-1999; 99WO-US22226.

XX 28-SEP-1998; 98US-0102161.

XX 28-SEP-1998; 98US-0102180.

XX 29-SEP-1998; 98US-0102380.

XX 08-OCT-1998; 98US-0103845.

XX 27-OCT-1998; 98US-0105877.

XX (CHIR) CHIRON CORP.

XX (HYSE-) HYSEQ INC.

XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;

XX Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;

XX Lamson G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;

XX Leshkowitz D, Kita D, Garcia V, Jones LW, Stache-Crain B;

XX WPI; 2000-293155/25.

XX Polynucleotide library comprising 1079 defined sequences, useful in

XX the form of an array to detect cancer or susceptibility to cancer -

XX Claim 1; Page 333; 502pp; English.

XX The present invention describes a library of polynucleotides comprising

XX 1079 nucleotide sequences (given in AAH30067 to AAH31145). Also described

XX are: (1) an isolated polynucleotide (I) having at least 90% identity to

XX one of the 1079 sequences; (2) a recombinant host cell containing (I);

XX (3) an isolated polypeptide (II) encoded by (I); (4) an antibody that

XX specifically binds to (II); (5) a vector comprising (I); and (6) a method

XX of detecting differentially expressed genes correlated with a cancerous

XX state of a mammalian cell comprising detecting a gene product encoded by

XX 65 of the 1079 sequences given in the specification. The polynucleotides

XX are used to monitor patients having (or susceptible) to cancer to detect

XX potentially malignant events at a molecular level before they are

XX detectable at a gross morphological level. The polynucleotides are also

XX useful for monitoring the efficacy of various therapies and preventive

XX interventions. Polynucleotide probes based on the disclosed sequences

XX are useful for chromosome mapping and detection of transcription levels.

XX The 1079 polynucleotide sequences were derived from a human colon cancer

XX cell line Km12L4-A cDNA library.

SQ Sequence 408 BP; 117 A; 96 C; 98 G; 97 T; 0 other;

Alignment Scores:

Pred. No.: 4,22e-62 Length: 408
Score: 717.00 Matches: 134
Percent Similarity: 99.26% Conservative: 0
Best Local Similarity: 99.26% Mismatches: 1
Query Match: 33.66% Indels: 0
DB: 21 Gaps: 0


```
Db 115 -----TTCCGCGAGGGGACCGACGCTACGCGCGGGGAGTGGCCCGGGGTGGTCTCTG 168
Qy 66 TyrLeuGluIleSerLeuArgLeuHisArgLeuLeuArgAspSerGluAlaPheCysHis 85
Db 169 AGCATGGAACGGGCGCTGGCTCCCGGGGAGCCCTCCGGCGCTTCGGCTGGCTGGCGC 228
Qy 86 ArgAsnCysSerAla-----AlaProGlnProGlu 95
Db 229 ACCAGTGTGGCGCGACATTCCTCGTGGAGCTGGACCCCGAGCTGGTCCCGACCGCGGCC 288
Qy 96 ProAlaAlaGlyLeuAlaSerTyrProGluLeuArgLeuPheGlyGlyLeuArgArg 115
Db 289 CAGGCTCGGGCGCGCGCTCGGACCTGAGCTTTCGGGGGCTTCGGGTGCGC 348
Qy 116 AlaHisCysLeuArgCysGlnGlyLeuProAlaPheArgGlnSerGlnProSer 135
Db 349 GCTGCTGCTGCGCGCTGC---CTCGGGCGCGCGCGCCACCTCGCTC-----AGC 399
Qy 136 ArgGluValLeuAlaAspPheGlnArgGluProTyrLysPheLeuGlnPheAlaTyr 155
Db 400 GAAGAGATGAGCTGGAGTTCGCAAGCGAGCGCCCTCAACTACTGACGTCGCTTAC 459
Qy 156 PheLysAlaAsnAsnLeuProLysAlaIleAlaAlaHisThrPheLeuLeuLysHis 175
Db 460 TTCAAGATCAACAAGTTGAGAAAGCTGTGTGTCAGCACACACCTTCTCGTGGCAAT 519
Qy 176 ProAspAspGluMetLysArgAsnMetAlaTyrTyrLysSerLeuProGlyAlaGlu 195
Db 520 CTTGAGCATCATGGAATGAGCAGAACCTAGACTATTACCAACCATCTCTGGAGTGAAG 579
Qy 196 AspTyr---IleLysAspLeuGluThrLysSerTyrGluSerLeuPheIleArgAlaVal 214
Db 580 GAGCGCGACTTCAGGATCTTGAGACTCAACCCCATATGCAAGATTTTCGACTGGAGTG 639
Qy 215 ArgAlaTyrAsnGlyGluAsnTrpArgThrSerIleThrAspMetGluLeuAlaLeuPro 234
Db 640 CGACTCTACTAGAGAAACAGCACAGGAAAGCTGTGCCCCACCTAGAGCGCGCTGCA 699
Qy 235 AspPheLeuAlaPheTyrGluCysLeuAlaAlaCysGluGlySerArgGluLeuLys 254
Db 700 GAATACTTTGTGGCTATAGAGTGCCTGCGTCCCTCTCGAAGGCGCTTATGACTACGAT 759
Qy 255 AspPheLys-----AspPheTyrLeuSerIleAlaAspHisTyrVal 268
Db 760 GGTACAACTACTTGTAGTACAACTGACCTTCTCCAGGCCATCACAGATCATATACATC 819
Qy 269 GluValLeuGluCysLysIleGlnCysGluGluAsnLeu---ThrProValIleGly 287
Db 820 CAGGCTCTCACTGTAAGCAGAACTGTGTACGGAGCTGTCTCCACCCCAAGTCGAGAG 879
Qy 288 TyrProValGluLysPheValAlaThrMetTyrHisTyrLeuGlnPheAlaTyrTyrLys 307
Db 880 AAGCCCTTTGAGAGCTTCTCCCATCGCATATATATCTGCACTTTGCTACTATATAC 939
Qy 308 LeuAsnAspLeuLysAsnAlaProCysAlaValSerTyrLeuLeuPheAspGlnAsn 327
Db 940 AITGGGAATTATACAGAGCTGTGAAATGTGCAAGACCTATCTTCTTCTTCCCAAT 999
Qy 328 AspLysValMetGluGlnAsnLeuValTyrTyrGlnTyrHisArgAspThrTrpGlyLeu 347
Db 1000 GACGAGGTGATGAACCAAAATTTGGCTATTATGACGCTATG-----CTT 1044
Qy 348 SerAspGluHisPheGln-----ProArgProGluAlaValGlnPheAsnVal 364
Db 1045 GGAGAAGAACACACACAGATCCATCGGCCCGCTGAGAGTGGCAAGAGTACCGACAGCGA 1104
Qy 365 ThrThrLeuGlyGluLeuTyrAspPheAlaLysGlu-----AsnIleMet 380
Db 1105 AGCCTACTGAAAGAACTGTTCTTCTCGCTTATGATGTTTTTGGAAATTCCTTTGTG 1164
Qy 381 AspAspAspGluGlyGluValValGluTyrValAspAspLeuLeuGluLeuGluThr 400
Db 1165 GATCCGGATTTCAGACTCCAGAGAGTGTATCCACAGAGATTCCAGAGAAACAGAG 1224
```

Qy 401 Ser 401
Db 1225 TCA 1227

RESULT 12

AAH14513
ID AAH14513 standard; cDNA; 2524 BP.

XX AC AAH14513;

XX 26-JUN-2001 (first entry)

DE Human cDNA sequence SEQ ID NO:12044.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

XX EPI074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

XX 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai X, Otsuki T;

XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602

XX full-length cDNAs defined in the specification, and for the detection

XX of or diagnosis of the abnormality of the proteins encoded by the

XX full-length cDNAs.

XX Claim 8; SEQ ID 12044; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602

XX full-length cDNAs defined in the specification. Where a primer set

XX comprises: (a) an oligo-dr primer and an oligonucleotide complementary

XX to the complementary strand of a polynucleotide which comprises one of

XX the 5602 nucleotide sequences defined in the specification, where the

XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination

XX of an oligonucleotide comprising a sequence complementary to the

XX complementary strand of a polynucleotide which comprises a 5'-end

XX sequence and an oligonucleotide comprising a sequence complementary to a

XX polynucleotide which comprises a 3'-end sequence, where the

XX oligonucleotide comprises at least 15 nucleotides and the combination

XX of the 5'-end sequence/3'-end sequence is selected from those defined in

XX the specification. The primer sets can be used in antisense therapy and

XX in gene therapy. The primers are useful for synthesizing polynucleotides,

XX particularly full-length cDNAs. The primers are also useful for the

XX detection and/or diagnosis of the abnormality of the proteins encoded by

XX the full-length cDNAs. The primers allow obtaining of the full-length

XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and

XX AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to

XX AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632

XX represent oligonucleotides, all of which are used in the exemplification

XX of the present invention.

XX SQ Sequence 2524 BP; 583 A; 747 C; 704 G; 490 T; 0 other;

Alignment Scores:

Pred. No.: 1.67e-46 Length: 2524

Score: 571.00 Matches: 145

Percent Similarity: 50.59% Conservative: 68
Best Local Similarity: 34.44% Mismatches: 168
Query Match: 26.81% Indels: 40
DB: 22 Gaps: 11

US-09-729-674-2 (1-401) x AAH14513 (1-2524)

QY	6	ArgGlyAlaAlaAlaLeuAlaLeuLeuCysValAlaCysAlaLeuArgAlaGlyArg	25	
Db	25	CGCGCGTTGAAGCTGTCACCACTGCTGCTGCTGCGCGCT-----GCTCCCAA	78	
QY	26	AlaGlnTyrGluArgTyr-SerPheArgSerPheProArgaspGluLeuMetProLeuGlu	45	
Db	79	GCAGAGTCGAGTCCAGGACGATGGGGCATGGTACGCTGATCTGCTC-----	129	
QY	46	SerAlaTyrArgHisAlaLeuAspLysTyrSerGlyGluHisThrAlaGluSer***Gly	65	
Db	130	-----TTCCGCGAGGGACCCAGCACTACCGCGCGGGGACTGGCGCGGGTGCCTG	183	
QY	66	TyrLeuGluHisSerLeuArgLeuHisArgLeuLeuArgaspSerGluAlaPheCysHis	85	
Db	184	AGCATGGAACGGGCGCTGCGCTCCCGGCGACCTTCGCGCCCTTCGCTGCGCTGCGC	243	
QY	86	ArgAsnCysSerAla-----	AlaProGlnProGlu	95
Db	244	ACCCAGTGTGGCGCGACTTCCTCGTGGGAGCTGGACCCGAGTGGCCCCCAGCGCGC	303	
QY	96	ProAlaAlaGlyLeuAlaSerTyrProGluLeuArgLeuPheGlyGlyLeuLeuArgArg	115	
Db	304	CAGGCTTCGGGCGCGCGCGCTCGCGACCTGAGCTTCTTCGGGGGCGCTTCTGCGTCG	363	
QY	116	AlaHisCysLeuLysArgCysLysGlnGlyLeuProAlaPheArgGlnSerGlnProSer	135	
Db	364	GCTGCTGCTCGCGCGCTGC---CTCGGCGCGCGCGCGCCACTCGCTC-----	AGC	414
QY	136	ArgGluValLeuAlaAspPheGlnArgGluProTyrLysPheLeuGlnPheAlaTyr	155	
Db	415	GAAGAGATGGAGTCGGAGTTCGCAAGCGGAGCCCTACAACTACCTCGAGTCGCGCTAC	474	
QY	156	PheLysAlaAsnAsnLeuProLysAlaIleAlaAlaHisThrPheLeuLeuLysHis	175	
Db	475	TTCAAGATCAAACAAGTTGGAAAGCTGTCTCGACGACACACCTTCTTCGTGGGCAAT	534	
QY	176	ProAspAspGluMetMetLysArgAsnMetAlaTyrTyrLysSerLeuProGlyValaGlu	195	
Db	535	CCTGAGCATGAGAAATGAGCAGAACCTAGACTATTACCAAAACCATGTCTGGAGTGAAG	594	
QY	196	AspTyr---IleLysAspLeuGluThrLysSerTyrGluSerLeuPheIleArgAlaVal	214	
Db	595	GAGCGCGACTTCAGAGATCTTGAGACTCAACCCCATATGCAAGAAATTCGACTGGGAGTG	654	
QY	215	ArgAlaTyrAsnGlyGluAsnTrpArgThrSerIleThrAspMetGluLeuAlaLeuPro	234	
Db	655	CGACTCTACTCAGAGAAACAGCACAGGAAGCTGTGCCCCACCTAGAGGGCGGCTGCAA	714	
QY	235	AspPhePheLysAlaPheTyrGluCysLeuAlaAlaCysGluGlySerArgGluIleLys	254	
Db	715	GAATACTTTGTGCGCTATGAGAGTGCCTGCCCTCTCGAAGGGCCCTATGACTACGAT	774	
QY	255	AspPheLys-----AspPheTyrLeuSerIleAlaAspHisTyrVal	268	
Db	775	GGGTACAACACTCTTGAGTACAACTGACCTTCTCCAGGCCATCACAGATCATTACATC	834	
QY	269	GluValLeuGluCysLysIleGlnCysGluGluAsnLeu---ThrProValIleGlyGly	287	
Db	835	CAGGTCTCTCAACTGTAAAGCAGAACTGTGTCAAGGAGCTTGCTGCCAACCCCAAGTCGAGAG	894	
QY	288	TyrProValGluLysPheValAlaThrMetTyrHisTyrLeuGlnPheAlaTyrTyrLys	307	
Db	895	AAGCCCTTTGAGACTTCTCCCTCGCATTAATATATCTGCAGTTTGGCTTACTTAAC	954	
QY	308	LeuAsnAspLeuLysAsnAlaAlaProCysAlaValSerTyrLeuLeuPheAspGlnAsn	327	

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 30, 2004, 03:41:40 ; Search time 4828 Seconds
(without alignments)
3397.838 Million cell updates/sec

Title: US-09-729-674-2
Perfect score: 2130
Sequence: 1 MEPGRGAALLALLCVACA.....DDEGEVVEYVDDLELEETS 401

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ _p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US09729674/runat_29012004_095314_12887/app_query.fasta_1.583
-DB=GenEmbl -QFMT=fastcap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -NATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=75 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09729674@cgn_1_3508 @runat_29012004_095314_12887 -NCFU=6 -ICPU=3
-NO MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1:	gb.ba.*	AC112211 Homo sapi
2:	gb.hcg.*	AC112211 Homo sapi
3:	gb.in.*	AC136358 Homo sapi
4:	gb.em.*	AC122652 Rattus no
5:	gb.ov.*	AC032013 Mus muscu
6:	gb.pat.*	AL590968 Mouse DNA
7:	gb.ph.*	AC112375 Rattus no
8:	gb.pl.*	AC113636 Rattus no
9:	gb.pr.*	AC091172 Homo sapi
10:	gb.ro.*	AC012192 Homo sapi
11:	gb.sts.*	BD158280 Primer fo
12:	gb.sy.*	AK027648 Homo sapi
13:	gb.in.*	BD156505 Primer fo
14:	gb.vi.*	AK027680 Homo sapi
15:	em.ba.*	AX136191 Sequence
16:	em.fun.*	BD123548 Secretary
17:	em.hum.*	AK075418 Homo sapi
18:	em.in.*	AX704765 Sequence
19:	em.mu.*	BD156581 Primer fo
20:	em.om.*	AK027697 Homo sapi
21:	em.or.*	AF097432 Homo sapi
22:	em.ov.*	AF456412 Mus muscu
23:	em.pat.*	BC024047 Mus muscu
24:	em.pl.*	
25:	em.pl.*	
26:	em.ro.*	
27:	em.sts.*	
28:	em.un.*	

ALIGNMENTS

RESULT 1

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Length	DB	ID	Description
1	2128	99.9	2307	9	HSAJ6470	AJ006470 Homo sapi
2	2128	99.9	3871	6	BD140451	BD140451 Secreted
3	2125	99.8	1439	9	BC008745	BC008745 Homo sapi
4	2121	99.6	1961	6	BD156517	BD156517 Primer fo
5	2121	99.6	1961	9	AK001634	AK001634 Homo sapi
6	2101	98.6	2010	9	AK075401	AK075401 Homo sapi
7	2049.5	96.2	1922	9	AK091772	AK091772 Homo sapi
8	1986.5	93.3	1903	9	AK092990	AK092990 Homo sapi
9	1894	88.9	1685	10	MMAJ8469	AU008469 Mus muscu
10	1893	88.9	1663	10	BC049890	BC049890 Mus muscu
11	1827.5	85.8	1810	9	AK097650	AK097650 Homo sapi
12	1190.5	55.9	1665	5	GGCARPR	X97607 G.gallus mr
13	1155	54.2	818	6	BD148800	BD148800 Primer fo
14	1103.5	51.8	1903	10	BC031856	BC031856 Mus muscu
15	1090.5	51.2	2079	9	BSA250583	AJ750583 Homo sapi
16	1089.5	51.2	1889	9	BC001047	BC001047 Homo sapi
17	1089.5	51.2	2037	9	BC007942	BC007942 Homo sapi
18	1089.5	51.2	2443	6	AX714040	AX714040 Sequence
19	1089.5	51.2	2443	9	AK056085	AK056085 Homo sapi
20	1088.5	51.1	1745	9	BC011701	BC011701 Homo sapi
21	1088.5	51.1	2347	6	AX335490	AX335490 Sequence
22	1088.5	51.1	2347	9	HSU47821	XU47821 Homo sapien
23	1027	48.2	1407	10	RNSC65MR	X65454 R.norvegicu
24	831	39.0	182509	9	AC112211	AC112211 Homo sapi
25	791	37.1	64706	2	AC136358	AC136358 Homo sapi
26	675.5	31.7	226615	2	AC122652	AC122652 Rattus no
27	627.5	29.5	176789	2	AC032013	AC032013 Mus muscu
28	627.5	29.5	217092	10	AL590968	AL590968 Mouse DNA
29	611	28.7	226822	2	AC112375	AC112375 Rattus no
30	611	28.7	241607	2	AC113636	AC113636 Rattus no
31	591	27.7	178345	9	AC091172	AC091172 Homo sapi
32	591	27.7	196735	2	AC012192	AC012192 Homo sapi
33	572.5	26.9	2993	6	BD158280	BD158280 Primer fo
34	572.5	26.9	2993	9	AK027648	AK027648 Homo sapi
35	571	26.8	2524	6	BD156505	BD156505 Primer fo
36	571	26.8	2524	9	AK027680	AK027680 Homo sapi
37	571	26.8	2563	6	AX136191	AX136191 Sequence
38	571	26.8	2563	6	BD123548	BD123548 Secretary
39	571	26.8	2563	9	AK075418	AK075418 Homo sapi
40	571	26.8	2583	6	AX704765	AX704765 Sequence
41	571	26.8	2585	6	BD156581	BD156581 Primer fo
42	571	26.8	2585	9	AK027697	AK027697 Homo sapi
43	567	26.5	2800	9	AF097432	AF097432 Homo sapi
44	565.5	26.5	2880	10	AF456412	AF456412 Mus muscu
45	545	25.6	2569	10	BC024047	BC024047 Mus muscu


```

Db      1056 CACAGGACACGTGGGGCTCTCAGATGAGCACTTCCAGCCAGACCTGAAGCAGTTCAG 1115
Qy      361 PhePheAsnValThrThreuglnlysluLeuTyArgPheAlaLysGluAsnIleMet 380
Db      1116 TTCTTAATGTGACCACACTCCAGAGGAGCTGTATGACTTTGCTAAGGAAAATATATG 1175
Qy      381 AspAspAspGluGlyValValGluTyArgValAspAspLeuLeuGluGluThr 400
Db      1176 GATGATGATGAGGAGAAATTTGGATATGTGGATGACCTTTGGAACTGGAGGAGACC 1235
Qy      401 Ser 401
Db      1236 AGC 1238

RESULT 4
BD156517
LOCUS      BD156517              1961 bp      DNA      linear      PAT 17-JAN-2003
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION  BD156517
VERSION    BD156517.1 GI:27862275
KEYWORDS  JP 2002191363-A/11360.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 1961)
AUTHORS   Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
            Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
TITLE     Primer for synthesizing full-length cDNA and use thereof
JOURNAL   Patent: JP 2002191363-A 11360 09-JUL-2002;
            HELIX RESEARCH INSTITUTE
COMMENT   OS Homo sapiens (human)
            PN JP 2002191363-A/11360
            PD 09-JUL-2002
            PF 28-JUL-2000 JP 2000280990
            PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
            PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
            PI KEIICHI NAGAI, TETSUJI OTSUKI
            PC
            C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
            10, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
            Primer for synthesizing full-length cDNA and use thereof PH Key
            Location/Qualifiers
            FT CDS      Location/Qualifiers
            1..1961
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            417 a 579 c 481 g 484 t
            BASE COUNT  417 a 579 c 481 g 484 t
            ORIGIN

Alignment Scores:
Pred. No.:      2,8e-176      Length:      1961
Score:          2121.00      Matches:      399
Percent Similarity: 99.50%      Conservative: 0
Best Local Similarity: 99.50%      Mismatches: 2
Query Match:    99.58%      Indels:      0
DB:             6           Gaps:        0

US-09-729-674-2 (1-401) x BD156517 (1-1961)
Qy      1 MetGluProGlyArgArgGlyValAlaAlaLeuLeuAlaLeuLeuCysValAlaCysAla 20
Db      40 ATGAGCGGGGCGCGGGGGCGCGCGCTGCTAGCTGTGCTGCGGCTCGCGG 99
Qy      21 LeuArgAlaGlyArgAlaGlnTyArgTySerPheArgSerPheProArgAspGlu 40
Db      100 CTGGCGCGGGCGCGCGCCCAATACGAACGCTACAGCTTCCGAGCTTCCACGGGACGAG 159
Qy      41 LeuMetProLeuGluSerAlaTyArgHisAlaLeuAspLysTySerGlyGluHisTrp 60

```

```

Db      160 CTGATCCGCTCAGTCCGCTACCGGCACGCGCTGGCAAGTACAGCGCGCAGCACTGG 219
Qy      61 AlaGluSer***GlyTyLeuGluLleSerLeuAtcLeuHisArgLeuLeuArgAspSer 80
Db      220 GCCGAGAGCGTGGGCTTACCTGGAGATCAGCGCTGCGGCTCACCGCTTGTGTGGCGACG 279
Qy      81 GluAlaPheCysHisArgAsnCysSerAlaAlaProGlnProGluProAlaAlaGlyLeu 100
Db      280 GAGGCTTCTGCGCACCGCAACTGCAGCGCGCGCGCGCAGCGCGCGCGCGCGCGCTC 339
Qy      101 AlaSerTyProGluLeuArgLeuPheGlyGlyLeuLeuArgArgAlaHisCysLeuLys 120
Db      340 GCCAGCTATCCGAGCTGCGCTCTTTCGGGGGCGCTGCTCGCGCGCGCGCTGCTCAAG 399
Qy      121 ArgCysLysGlnGlyLeuProAlaPheArgGlnSerGlnProSerArgGluValLeuAla 140
Db      400 CGCTGCAAGCAGGCGCTGCGAGCCTTCCGCCAGTCCCGAGCCCGAGCGCGAGGTGGCG 459
Qy      141 AspPheGlnArgArgGluProTyLysPheLeuGlnPheAlaTyPheLysAlaAsnAsn 160
Db      460 GACTTCCAGCGCGCGAGCGCTTACAGGTTCCTGCGAGTTCGCTTACTTCAAGCAATAAT 519
Qy      161 LeuProLysAlaAlaLeuAlaHisThrPheLeuLeuLysHisProAspAspGluMet 180
Db      520 CTTCCCAAAGCCATCGCCGCTGCTCACCTTCTTACTGAAGCATCCTGATCGAAATG 579
Qy      181 MetLysArgAsnMetAlaTyTyLysSerLeuProGlyValaGluAspTyLysLeuAsp 200
Db      580 ATGAAGAGGAACATGTCATATTATAAGAGCTGCTGCTGGTGGCGAGGACTTACATTAAAGAC 639
Qy      201 LeuGluThrLysSerTyGluSerLeuPheLeuPheLeuAlaValArgAlaTyrsnglyGlu 220
Db      640 CTGGAACCAAGTCATATGAAGCCTGTTCATCCGAGCAGTCCGGGCATACAAACGGTGAG 699
Qy      221 AsnTrpArgThrSerIleThrAspMetGluLeuAlaLeuProAspPhePheLysAlaPhe 240
Db      700 AACTGGAGAACATCCATCACAGCATGGAGCTGGGCTTCCGACTTCTTCAAGCCCTTT 759
Qy      241 TyGluCysLeuAlaAlaCysGluGlySerArgGluLleLysAspPheLysAspPheTy 260
Db      760 TACGAGTGTCTCGAGCGCTCGAGGCTTCCAGGGAGATCAAGGACTTCAAGGATTCTTAC 819
Qy      261 LeuSerIleAlaAspHisTyValGluValLeuGluCysLysLleGlnCysGluGluAsn 280
Db      820 CTTTCCATAGCAGCATATTGTAGAAGTCTTGGAAATCCAAATACAGTGTGAAGAGAAC 879
Qy      281 LeuThrProValIleGlyTyProValGluLysPheValAlaThrMetTyHisTy 300
Db      880 CTCACCCAGTTTATAGGAGGCTATCCGCTTGAGAAATTTGTGGCTACCATGTATCATTAC 939
Qy      301 LeuGlnPheAlaTyTyLysLeuAsnAspLeuLysAsnAlaAlaProCysAlaValSer 320
Db      940 TTGCAGTTTGCTTATTAAAGTTGAACGCTGGAAGATGAGCCCGCTGTGCGATCAGC 999
Qy      321 TyLeuLeuPheAspGlnAsnAspLysValMetGlnGlnAsnLeuValTyTyGlnTy 340
Db      1000 TATCTGCTTTTGATCAAGATGACAGGCTCATGCGAGCAGACCTTGGTGTATTACAGTAC 1059
Qy      341 HisArgAspThrTrpGlyLeuSerAspGluHisPheGlnProArgProGluAlaValGln 360
Db      1060 CACAGGACAGCTGGGGGCTCTCAGATGAGCACTTCCAGCCCGAGCCTTGAAGCAGTTCAG 1119
Qy      361 PhePheAsnValThrThrLeuGlnLysGluLeuTyArgPheAlaLysGluAsnIleMet 380
Db      1120 TTTCTTAATGTGACCACTCCAGAGGAGCTGTATGACTTTGTAGGAAAATATATATG 1179
Qy      381 AspAspAspGluGlyGluValValGluTyValAspAspLeuLeuGluGluThr 400
Db      1180 GATGATGATGAGGAGAGAGTTGTGGAATATGTGATGACCTTCTTGGAACTTGGAGAGACC 1239
Qy      401 Ser 401

```


101 AlaserTyrProGluLeuArgGluPheGlyGlyLeuLeuArgArgAlaHisCysLeuLys 120

ORGANIS

Homo sapiens


```

Db      684  AACTGGAGAACATCCATCACACAGATGAGCTGGCCCTCCGACTTCTTCAAGCCCTTT 743
QY      241  TyrGluCysLeuAlaAAlaCysGluGlySerArgGluLeuLysAspPheLysAspPheTyr 260
Db      744  TACGAGTGTCTCGCAGCCTCGCAGGTTCCAGGAGATCAAGGACTTCAAGGATTTCTAC 803
QY      261  LeuSerIleAlaAspHisTyrValGluValLeuGluCysLysIleGlnCysGluGluAsn 280
Db      804  CTTTCCATAGCA----- 815
QY      281  LeuThrProValIleGlyGlyTyrProValGluLysPheValAlaThrMetTyrHisTyr 300
Db      815  ----- 815
QY      301  LeuGlnPheAlaTyrTyrLysLeuAsnAspLeuLysAsnAlaAlaProCysAlaValSer 320
Db      816  -----GTGAACACCTGAGAAATGACGCCCTGTGCAGTCAGC 854
QY      321  TyrLeuLeuPheAspGlnAsnAspLysValMetGlnGlnAsnLeuValTyrTyrGlnTyr 340
Db      855  TATCTGCTCTTATCAGATGACAGAGCTCATGCGACAGAACCTGGTGATTACAGTAC 914
QY      341  HisArgAspThrTyrGlyLeuSerAspGluHisPheGlnProArgProGluAlaValGln 360
Db      915  CACAGGACACGCTGGGGCTCTCAGATGAGCACTTCAGGCCACAGACCTGAAGCAGTTTCA 974
QY      361  PhePheAsnValThrThrLeuGlnLysGluLeuTyrAspPheAlaLysGluAsnIleMet 380
Db      975  TCTTTATATGACACACTCAGAGAGAGCTGTATGACTTTGTCTAAGGAAATATATATG 1034
QY      381  AspAspGluGlyGluValValGluTyrValAspAspLeuLeuGluGluThr 400
Db      1035  GATGATGATGAGGAGAAATGTGGAATATGTGATGACCTTTTGAACCTGGAGAGACC 1094
QY      401  Ser 401
Db      1095  AGC 1097

RESULT 12
LOCUS   GGCARPR 1665 bp mRNA linear VRT 26-JUN-1997
DEFINITION G.gallus mRNA for cartilage associated protein.
ACCESSION X97607
VERSION X97607.1 GI:1296525
KEYWORDS cartilage associated protein.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
REFERENCE 1
AUTHORS Castagnola, P., Gennari, M., Morello, R., Tonachini, L., Marin, O.,
Gaggero, A. and Cancedda, R.
TITLE Cartilage associated protein (CASP) is a novel developmentally
regulated chick embryo protein
J Cell. Sci. 110 (Pt 12), 1351-1359 (1997)
PUBMED 97360293
REFERENCE 2 (bases 1 to 1665)
AUTHORS Castagnola, P.
DIRECT SUBMISSION
SUBMITTED (29-APR-1996) P. Castagnola, Centro di Biotecnologie
Avanzate, Differenziamento Cellulare, L.go R. Benzi, 10 Genova,
16132, ITALY
FEATURES
Location/Qualifiers
1..1665
/organism="Gallus gallus"
/mol_type="mRNA"
/db_xref="taxon:9031"
/cell_type="chondrocyte"
/dev_stage="embryo"
/lab_host="E.coli"
43..858
CDS

```

```

/codon_start=1
/product="cartilage associated protein"
/protein_id="CAA66206.1"
/db_xref="GI:1296526"
/db_xref="SMISS:PR0T:Q08030"
/translation="NMRLTLLAALLATAGAQYERYFRFPFRLDMLPESAYRYGLDQY
STENPESVYLEVSRMLRRLURDTEAFCHNCSSAGPLTAPPDAGELEAURLLAGV
LRRACLRCKQGLFAFQAQFRELLEEFORREPKYKLFQFAFYKXANLPKAAAH
FLKHPDEMORMNWAYKSIDABEHLKDLREKVENLFRVAYNGNWRISID
MELALPDFDKTYDDGCIACGEGSRGDKDFKDFYLSIADHYIEVLACKVQFDIT"
BASE COUNT 470 a 389 c 388 g 418 t
ORIGIN
Alignment Scores:
Pred. No.: 8,16e-95 Length: 1665
Score: 1190.50 Matches: 238
Percent Similarity: 68.15% Conservative: 38
Best Local Similarity: 58.77% Mismatches: 36
Query Match: 55.89% Indels: 94
DB: 5 Gaps: 6
US-09-729-674-2 (1-401) x GGCARPR (1-1665)
QY 3 ProGlyArgArgGlyAla-----AlaAlaLeuLeuAlaLeuLeuCysVal 17
Db 4 CCACGGCGACGTGGTGGCGCGCGCGCGGAGGATGTGGCGGACGCTGTGGCG 63
QY 18 AlaCysAlaLeuArgAlaGlyArgAlaGlnTyrGluArgTyrSerPheArgSerPhePro 37
Db 64 GCGCTTCTGGCGACGCGCGGC---GCGCAGTACGAGCGGTACAGCTTCGCGAGCTCCCG 120
QY 38 ArgAspGluLeuMetProLeuGluSerAlaTyrArgHisAlaLeuAspLysTyrSerGly 57
Db 121 CGGACGAGCTGATGCCCTGGAATCCGCTACCGCTACGCGCTGGACGACGACGACG 180
QY 58 GluHisTrpAlaGluSer***GlyTyrLeuGluIleSerLeuArgLeuHisArgLeuLeu 77
Db 181 GAGAACTGCCCGAGAGCGTCAGCTACCTGGAGGTGAGCATCGCGCTGTACCGCTGCTG 240
QY 78 ArgAspSerGluAlaPheCysHisArgAsnCysSerAlaAla-----ProGlnPro 94
Db 241 CGCGACACGAGGCGCTTTCGCCACCACTGCGAGCTCCGCGCGCGCGCTCACCGCCCG 300
QY 95 GluProAlaAlaGlyLeuAlaSerTyrProGluLeuArgLeuPheGlyGlyLeuLeuArg 114
Db 301 CGCGCGCGCGATGGG-----CAGCTGCCGAGCTGCCCTGCTGGCGGAGTGTGGCG 354
QY 115 ArgAlaHisCysLeuLysArgCysLysGlnGlyLeuProAlaPheArgGlnSerGlnPro 134
Db 355 CGAGCGCAATGCTCGCGCGTGAAGCAGGCGGTGGCGCGCTTCGCGCGCGCGCGAGCC 414
QY 135 SerArgGluValLeuAlaAspPheGlnArgGluProTyrLysPheLeuGlnPheAla 154
Db 415 GCGCGCGAGCTGCTCGAGAGTTCCAGCGCGCGAGAGCCCTCAAGTACTCTCAGTTCCGC 474
QY 155 TyrPheLysAlaAsnAsnLeuProLysAlaIleAlaAlaHisThrPheLeuLeuLys 174
Db 475 TACTTCAAGGCTATAAAGCTTCCAAAGCTATTGCGAGCGCTCACACATTTCTTCTGAAA 534
QY 175 HisProAspAspGluMetMetLysArgAsnMetAlaTyrTyrLysSerLeuProGlyAla 194
Db 535 CATCCAGATGATGAATGATCAAGAAATATGCGCTTACTATAAGAGATACCTGATGCT 594
QY 195 GluAspTyrIleLysAspLeuGluThrLysSerTyrGluSerLeuPheIleArgAlaVal 214
Db 595 GAGGAGCATTAAAGACTTGGAGCAAGAGCTTACGAGATCTCTTGTGAGGCTGTG 654
QY 215 ArgAlaTyrAsnGlyGluAsnTrpArgThrSerIleThrAspMetGluLeuAlaLeuPro 234
Db 655 AGAGCATCAATGGTGTGACAACTGAGAGACATCCATCTCAGATATGGAACCTGCTTCT 714
QY 235 AspPhePheLysAlaPheTyrGluCysLeuAlaAlaCysGluGlySerArgGluLeuLys 254

```

```

Db 715 GATTTCCTCAAACTACAGTACTGTATAGCAGCGCTGTGAGGCTCCGAGAGATTAAA 774
QY 255 AspPheLysAspPheTyrLeuSerIleAlaAspHisTyrValGluValLeuGluCysLys 274
Db 775 GATTTCAGAGACTTCTATCTCTATTCGAGATCATATTATGAAGTCTTGCATGCATA 834
QY 275 IleGlnCysGluGluAsnLeuThrProValIleGlyTyrProValGluLysPheVal 294
Db 835 GTGCAG----- 840
QY 295 AlaThrMetTyrHisTyrLeuGlnPheAlaTyrTyrLysLeuAsnAspLeuLysAsnAla 314
Db 840 ----- 840
QY 315 AlaProCysAlaValSerTyrLeuLeuPheAspGlnAsnAspLysValMetGlnGlnAsn 334
Db 840 ----- 840
QY 335 LeuValTyrTyrGlnTyrHisArgAspThrTrpGlyLeuSerAspGluHisPheGlnPro 354
Db 841 -----ITCGA-TATCAC----- 851
QY 355 ArgProGluAlaValGlnPhePheAsnValThrLeuGlnLysGluLeuTyrAspPhe 374
Db 852 -----AACATAACCCACACTCCAGTTGGAAATGTATGAATTC 887
QY 375 AlaLysGluAsnIleMetAspAspAspGluGlyGluValValGluTyrValAspLeu 394
Db 888 GCAAGGAAACATCTATGATGATGATGAGGTGAAGTGTGGAATTCCTTCATGAGTTG 947
QY 395 LeuGluLeuGluGlu 399
Db 948 TTAGAAATGGAGAA 962

RESULT 13
BD148800
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

OS Homo sapiens (human)
PN JP 2002191363-A/3643
PD 09-JUL-2002
PF 28-JUL-2000 JP 2000280990
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
PI SAITO,
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
PI KEIICHI NAGAI, TETSUJI OTSUKI
PC
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/
10, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
FT source
FT Location/Qualifiers
1. 818
/organism="Homo sapiens (human)"
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 159 a 271 c 236 g 148 t 4 others

```

ORIGIN

Alignment Scores: 4,11e-92 Length: 818
 Pred. No.: 1155.00 Matches: 238
 Score: 1155.00 Conservative: 1
 Percent Similarity: 91.57% Mismatches: 15
 Best Local Similarity: 91.19% Indels: 8
 Query Match: 54.23% Gaps: 1
 DB: 6

US-09-729-674-2 (1-401) x BD148800 (1-818)

```

QY 1 MetGluProGlyArgArgGlyAlaAlaAlaLeuLeuAlaLeuLeuCysValAlaCysAla 20
Db 40 ATGGAGCGGGGGCGGGGGCGGGGGCGGGGGCGGGGGCGGGGGCGGGGGCGGGGGCGGG 99
QY 21 LeuArgAlaGluValArgAlaGlnTyrGluArgTyrSerPheArgSerPheProArgaspGlu 40
Db 100 CTGGCGCGGGGGCGGGGGCGGGGGCGGGGGCGGGGGCGGGGGCGGGGGCGGGGGCGGG 159
QY 41 LeuMetProLeuGluSerAlaTyrArgHisAlaLeuAspLysTyrSerGlyGluHisTrp 60
Db 160 CTGATCGCGCTCGAGTCGGCTACCGCGCAGCGCTGGCAAGTACAGCGCGGAGCACTGG 219
QY 61 AlaGluSer***GlyTyrLeuGluIleSerLeuArgLeuHisArgLeuLeuArgaspSer 80
Db 220 GCCGAGAGCGTGGGCTACCTGGAGATCAGCCCTCGCGCTGCACCGCTTGTTCGCGCAGC 279
QY 81 GluAlaPheCysHisArgAsnCyssSerAlaAlaProGlnProGluProAlaAlaGlyLeu 100
Db 280 GAGGCTTCTGCCACCGCAACTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTC 339
QY 101 AlaSerTyrProGluLeuArgLeuPheGlyGlyLeuLeuArgArgAlaHisCysLeuLys 120
Db 340 GCCAGCTATCCGAGCTGCGCTCTTCGGGGGCGCTTTCGCGCGCGCGCGCGCGCTCAAG 399
QY 121 ArgCysLysGlnGlyLeuProAlaPheArgGlnSerGlnProSerArgGluValLeuAla 140
Db 400 CGCTGCAAGCAGCGGCTTCCAGCCTTCCGCGAGTCCAGCCAGCGCGCGAGGTGTGGGG 459
QY 141 AspPheGlnArgArgGluProTyrLysPheLeuGlnPheAlaTyrPheLysAlaAsnAsn 160
Db 460 GACTTCAGCGCGCGCGAGCCCTACAAGTTCCTGCGAGTTCGTTACTTCAAGCAATAAT 519
QY 161 LeuProLysAlaIleAlaAlaHisThrPheLeuLeuLysHisProAspAspGluMet 180
Db 520 CTCGCCAAGGCCATCGCCCTGCTCACCTTTCTACTGAAGCATCCTGATGACGAAATG 579
QY 181 MetLysArgAsnMetAlaTyrTyrLysSerLeuProGlyValaGluAspTyrIleLysAsp 200
Db 580 ATGAGAGAGAAACATGGCATATTATAAGACCTGCT-GGTGCGCGAGGACTTACATTAAAGAC 638
QY 201 LeuGluThrLys-SerTyrGlu-SerLeuPheIleArgAlaValaGluAlaTyrAsnGlyG 220
Db 639 CTGGAAACCAAGTTCATATGAAGGCTTTTCATCCGANCAGTGGGGGATACAAACGGTG 698
QY 220 LuAsnTrpArgThrSerIleThrAspMetGluLeuAla--LeuProAspPhePhe-LysA 239
Db 699 AAAACTGGAGAACATTCATCACAGCATGGAAGTGGCCCTTTCCGAAATTTCTTCAAAG 758
QY 239 lapherTyrGlu-----CysLeuAlaAlaCysGluGlySerArgGluLysLys 254
Db 759 NCTTTTACGAATGGTCTTGCANCCCTGGCGAAGGGTTNCAAGGGAAGATCAAG 811

```

RESULT 14

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

BC031856 1903 bp mRNA linear ROD 22-APR-2003
 Mus musculus SC65 synaptonemal complex protein, mRNA (cDNA clone
 MGC:25554 IMAGE:3968941), complete cds.
 BC031856
 BC031856.1 GI:21594429
 MGC.
 Mus musculus (house mouse)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2004, 07:10:40 ; Search time 4381.42 Seconds
(without alignments)
11232.506 Million cell updates/sec

Title: US-09-729-674-1_COPY_63_1265
Perfect score: 1203
Sequence: 1 atgagccggggccggggg.....tggaaactggaggagaccagc 1203

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pt.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vl.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1202.2	99.9	2307	9 HSAJ6470	AJ006470 Homo sapi
2	1202.2	99.9	3871	6 BD140451	BD140451 Secreted
3	1197.4	99.5	1439	9 BC008745	BC008745 Homo sapi
4	1197.4	99.5	1361	6 BD156517	BD156517 Primer fo
5	1197.4	99.5	1961	9 AK001634	AK001634 Homo sapi
6	1187	98.7	2010	9 AK075401	AK075401 Homo sapi
7	1111.4	92.4	1922	9 AK091772	AK091772 Homo sapi
8	1076.6	89.5	1903	9 AK092990	AK092990 Homo sapi
9	923.8	76.8	1663	10 BC049890	BC049890 Mus muscu
10	923.8	76.8	1685	10 MWAJ6469	AJ006469 Mus muscu
11	919	76.4	1810	9 AK097650	AK097650 Homo sapi
12	656.6	54.6	818	6 BD148800	BD148800 Primer fo
13	496.2	41.2	1665	5 GGCARPR	X97607 G.gallus nr
14	475.6	39.5	182509	9 AC112211	AC112211 Homo sapi
15	467	38.8	1889	9 BC001047	BC001047 Homo sapi
16	467	38.8	2037	9 BC007942	BC007942 Homo sapi
17	467	38.8	2443	6 AX714040	AX714040 Sequence
18	467	38.8	2443	9 AK056085	AK056085 Homo sapi
19	466.6	38.8	1745	9 BC011701	BC011701 Homo sapi
20	465.8	38.7	2347	6 AX335490	AX335490 Sequence
21	465.8	38.7	2347	9 HSAJ6470	AJ006470 Homo sapi
22	465.4	38.7	2079	9 HSAJ250583	AJ250583 Homo sapi
23	448.2	37.3	64706	2 AC136358	AC136358 Homo sapi
24	428.2	35.6	1903	10 BC031856	BC031856 Mus muscu
25	403	33.5	1407	10 RNSC65MR	X65454 R.norvegicu
26	347.4	28.5	226615	2 AC122652	AC122652 Rattus no
27	205.2	17.1	178345	9 AC091172	AC091172 Homo sapi
28	205.2	17.1	196735	2 AC012192	AC012192 Homo sapi
29	179.2	14.9	176789	2 AC032013	AC032013 Mus muscu
30	179.2	14.9	217092	10 AL590968	AL590968 Mouse DNA
31	174.4	14.5	226822	2 AC112575	AC112575 Rattus no
32	174.4	14.5	241607	2 AC113636	AC113636 Rattus no
33	156.4	13.0	2524	6 BD156505	BD156505 Primer fo
34	156.4	13.0	2524	9 AK027680	AK027680 Homo sapi
35	156.4	13.0	2563	6 AX136191	AX136191 Sequence
36	156.4	13.0	2563	6 BD123548	BD123548 Secretary
37	156.4	13.0	2563	9 AK075418	AK075418 Homo sapi
38	156.4	13.0	2583	6 AX704765	AX704765 Sequence
39	156.4	13.0	2585	6 BD156581	BD156581 Primer fo
40	156.4	13.0	2993	6 AK027697	AK027697 Homo sapi
41	156.4	13.0	2993	6 BD158280	BD158280 Primer fo
42	156.4	13.0	2993	9 AK027648	AK027648 Homo sapi
43	156	13.0	2249	10 MMU430350	AJ430350 Mus muscu
44	154.8	12.9	2600	9 AF097432	AF097432 Homo sapi
45	154.8	12.9	2829	9 AF097431	AF097431 Homo sapi

ALIGNMENTS

RESULT 1
HSAJ6470 HSAJ6470 2307 bp mRNA linear PRI 23-MAR-2000
LOCUS Homo sapiens mRNA for cartilage-associated protein (CASP).
DEFINITION
ACCESSION AJ006470.1 GI:3687321
VERSION
KEYWORDS cartilage-associated protein; CASP.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Tonachini,L., Morello,R., Monticone,M., Skaug,J., Scherer,S.W.,
Cancedda,R. and Castagnola,P.
TITLE CDNA cloning, characterization and chromosome mapping of the gene

encoding human cartilage associated protein (CRTAP)
 Cytogenet. Cell Genet. 87 (3-4), 191-194 (1999)

JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

2 (bases 1 to 2307)

Castagnola P., Biotechnology in Oncologia,
 Submitted (12-JUN-1998) Castagnola P., Biotechnology in Oncologia,
 Ist. Naz.le per la Ricerca sul Cancro, L.G. R. Benzi 10 Genoa,
 16132, ITALY

FEATURES

source

Location/Qualifiers

1..2307 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="Contig of pCPH2 and I.M.A.G.E. cloneID 1090695"
 /tissue type="fetal brain"
 /dev stage="19-23 weeks-old"

gene

1..2307 /gene="casp"

CDS

12..1217 /gene="casp"

/codon_start=1

/product="cartilage-associated protein (CASP)"

/protein_id="CAA07054.1"

/db_xref="GI:3687322"

/db_xref="SPTREMBL:O75718"

/translation="MEPGRGAALLALLVACALRAGRAQYRYSRSPFDELMP
 ESAYHADKYGEHWASVGLIEISLRLLRDLSEAFCHRNCSAAPQPFAAGLAS
 YPELRLFGGLRAHCLKRCQGLPAFRSQSPREVLADFORRPYKFLQPAFYKANN
 LPKAAAHFTLLKHPDDMMKRNMAVYKSLPGAEYDKLETYSYSLFTRAVRAN
 GENWSTSITDELALPDEFKAFYECLACEGSRLEIKDFEYLSIADHYVFLSKIQ
 CENLTPVIGGTPVKEFVATMHIQLFAYIKLNDLNAAPCAVSILFLDQNDKVMQON
 LVYQIHRDTWGLSDEHFQPRPEAVQFFNVITLQKELYDFAKENIMDDDEGEVVEYVD
 DLLELEETS"

sig_peptide

12..89 /gene="casp"

mat_peptide

90..1214 /gene="casp"

/product="cartilage-associated protein (CASP)"

545 a 633 c 548 g 578 t 3 others

BASE COUNT

ORIGIN

Query Match 99.9%; Score 1202.2; DB 9; Length 2307;

Best Local Similarity 99.8%; Pred. No. 2.6e-182;

Matches 1201; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

1 ATGGAGCGGGCGCGGGCGCGCGCTGCTAGCGCTGCTGTGCTGCTGCGCG 60

12 ATGGAGCGGGCGCGGGCGCGCGCTGCTAGCGCTGCTGTGCTGCTGCGCG 71

61 CTGGCGCGGGCGCGCGCGCAATACGACGCTACGCTGCTGCTGCTGCTGCTG 120

72 CTGGCGCGGGCGCGCGCGCAATACGACGCTACGCTGCTGCTGCTGCTGCTG 131

121 CTGATCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180

132 CTGATCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 191

181 GCGGAGAGSTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240

192 GCGGAGAGSTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 251

241 GAGGCGCTTCTGCGACCGCACTGCGCGCGCGCGCGCGCGCGCGCGCGCTC 300

252 GAGGCGCTTCTGCGACCGCACTGCGCGCGCGCGCGCGCGCGCGCGCGCTC 311

301 GCGAGTATCCGAGCTGCGCTTCTGCGCGCGCGCGCGCGCGCGCGCGCTC 360

312 GCGAGTATCCGAGCTGCGCTTCTGCGCGCGCGCGCGCGCGCGCGCGCTC 371

361 CGCTGCAACGAGGCTGCGCGCTTCTGCGCGCGCGCGCGCGCGCGCGCTC 420

RESULT 2

BD140451

LOCUS

DEFINITION

SECRETED

PROTEIN

AND

POLYNUCLEOTIDES

ENCODING THEM.

BD140451

VERSION

BD140451.1

GI:23235396

KEYWORDS

JP 2002506611-A/1.

SOURCE

Homo sapiens

(human)

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

3871 bp DNA linear PAT 18-SEP-2002

Secreted proteins and polynucleotides encoding them.

BD140451

SECRETED

PROTEIN

AND

POLYNUCLEOTIDES

ENCODING THEM.

BD140451

VERSION

BD140451.1

GI:23235396

KEYWORDS

JP 2002506611-A/1.

SOURCE

Homo sapiens

(human)

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

1 (bases 1 to 3871)
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Jacobs K., McCoy J.M., Lavallie E.R., Racie L.A.C., Evans C.,
 Merberg D., Treacy M., Agostino M.J., Li R.J.S., Wong G.G.,

372 CGCTCAAGCAGGCGCTGCCAGCTTCCGCGAGTCCAGCCAGCCAGCGCGGTCTGGCG 431

421 GACTTCCAGCGCGCGAGCCCTACAGTTCCCTGAGTTCCTGCTTACTTCCAGGCAATAAT 480

432 GACTTCCAGCGCGCGAGCCCTACAGTTCCCTGAGTTCCTGCTTACTTCCAGGCAATAAT 491

481 CTCCCAAGCCATCCCGCTGCTCACCTTTCTACTGAAGCATCTTCTGATGCGAAATG 540

492 CTCCCAAGCCATCCCGCTGCTCACCTTTCTACTGAAGCATCTTCTGATGCGAAATG 551

541 ATGAAGAGGACATGGCATATTATAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600

552 ATGAAGAGGACATGGCATATTATAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 611

601 CTGGAACCAAGTCATATGAAGCTTGTTCATCCGAGAGTGGCGGCATACACCGGTGAG 660

612 CTGGAACCAAGTCATATGAAGCTTGTTCATCCGAGAGTGGCGGCATACACCGGTGAG 671

661 AACTGGAGAACATCCATCAGACATGGAGCTGGCCCTTCCGACTTCTTCAAGCCCTT 720

672 AACTGGAGAACATCCATCAGACATGGAGCTGGCCCTTCCGACTTCTTCAAGCCCTT 731

721 TAGAGTGTCTCCAGCTGCGAGGCTTCCAGGAGATCAAGGACATTCAGGATTTCTAC 780

732 TAGAGTGTCTCCAGCTGCGAGGCTTCCAGGAGATCAAGGACATTCAGGATTTCTAC 791

781 CTTTCCATAGCAGATCATTTATGAGTTCGGAATGCAAAATACAGTGTGAAGAGAAC 840

792 CTTTCCATAGCAGATCATTTATGAGTTCGGAATGCAAAATACAGTGTGAAGAGAAC 851

841 CTCACCCAGTTATAGGAGCTATCCGTTGAGAAATTTGTGCTACCATGATCATTTAC 900

852 CTCACCCAGTTATAGGAGCTATCCGTTGAGAAATTTGTGCTACCATGATCATTTAC 911

901 TTGAGTGTGCTTATTAAGTTGAACGACCTGGAAGTGCAGCCCTTGTGCGTTCAGC 960

912 TTGAGTGTGCTTATTAAGTTGAACGACCTGGAAGTGCAGCCCTTGTGCGTTCAGC 971

961 TATCTGCTTTCATCAGATGACAGTTCATGAGTGCAGAGAACCTGCTGTATTCACGATC 1020

972 TATCTGCTTTCATCAGATGACAGTTCATGAGTGCAGAGAACCTGCTGTATTCACGATC 1031

1021 CACAGGACACACTTGGGCGCTCTCGGATGAGCACTTCCAGCCAGAGCTTGAAGCAGTTCAG 1080

1032 CACAGGACACACTTGGGCGCTCTCGGATGAGCACTTCCAGCCAGAGCTTGAAGCAGTTCAG 1091

1081 TTCTTTAATGTGACACACTCCAGAGAGGCTGTATGATCTTCTGCTAAGGAAATATAATG 1140

1092 TTCTTTAATGTGACACACTCCAGAGAGGCTGTATGATCTTCTGCTAAGGAAATATAATG 1151

1141 GATGATGATGAGGAGAGTGTGGAAATATGAGTGCAGCTTCTTGAACCTGGAGGAGACC 1200

1152 GATGATGATGAGGAGAGTGTGGAAATATGAGTGCAGCTTCTTGAACCTGGAGGAGACC 1211

1201 AGC 1203

1212 AGC 1214

Clark, H.F. and Recheil, K.
 Secreted proteins and polynucleotides encoding them
 Patent: JP 200206611-A 1 05-MAR-2002;
 GENETICS INSTITUTE INC
 COMMENT OS Homo sapiens (human)
 PN JP 200206611-A/1
 PD 05-MAR-2002
 PF 24-NOV-1998 JP 2000522118
 PR 26-NOV-1997 US 607066804, 23-NOV-1998 US 09/197886 PI
 KENNETH JACOBS, JOHN M MCCOY, EDWARD R LAVALLIE, LISA A COLLINS PI
 RACIE,
 PI CHERYL EVANS, DAVID MERBERG, MAURICE TREACY, MICHAEL J AGOSTINO,
 PI ROBERT J STEININGER II, GORDON G WONG, HILARY F CLARK, KIM PI
 FECHTEL,
 PC C12N15/09, C07K14/00, C12N1/21, C12N5/10, C12P19/34, C12P21/02, PC
 C12Q1/68//
 PC A61P39/00, A61P35/00, A61P37/04, A61P37/06, C12N15/00, C12N5/00 CC
 Secreted proteins and polynucleotides encoding them FH Key
 Location/Qualifiers
 FT source 1..3871
 FT /organism="Homo sapiens (human)".
 FT Location/Qualifiers
 1..3871
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 965 a 913 c 1033 g 958 t 2 others
 ORIGIN

Query Match 99.9%; Score 1202.2; DB 6; Length 3871;
 Best Local Similarity 100.0%; Pred. No. 2.5e-182;
 Matches 1203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCGCGGCGCGCGCGCGCGCGCTGTAGCGCTGTGCTGCTGCTGCTGCTGCTGCG 60
 DB 63 ATGAGCGCGGCGCGCGCGCGCGCGCTGTAGCGCTGTGCTGCTGCTGCTGCTGCG 122
 QY 61 CTGCGCGCGGCGCGCGCGCGCGCGCTGTAGCGCTGTGCTGCTGCTGCTGCTGCG 120
 DB 123 CTGCGCGCGGCGCGCGCGCGCGCGCTGTAGCGCTGTGCTGCTGCTGCTGCTGCG 182
 QY 121 CTGATCGCGCTCGAGTCGGCTACCGGACCGCGCTGAGCAAGTACAGCGGCGAGCTGG 180
 DB 183 CTGATCGCGCTCGAGTCGGCTACCGGACCGCGCTGAGCAAGTACAGCGGCGAGCTGG 242
 QY 181 GCCGAGAGCTGKGGCTACTGAGATCAGCTGCGGCTGACCGCTGCTGCGGCGAGC 240
 DB 243 GCCGAGAGCTGKGGCTACTGAGATCAGCTGCGGCTGACCGCTGCTGCGGCGAGC 302
 QY 241 GAGGCTTCTGCCACCGCAAATGAGCGCGCGCGCGCGAGCGCGCGCGCGCGCTC 300
 DB 303 GAGGCTTCTGCCACCGCAAATGAGCGCGCGCGCGCGAGCGCGCGCGCGCGCTC 362
 QY 301 GCCAGCTATCCGAGCTCGCTCTTGGGGGCTGCTGCGCGCGCTGCTGCGCGCTCAAG 360
 DB 363 SCCAGCTATCCGAGCTCGCTCTTGGGGGCTGCTGCGCGCGCTGCTGCGCGCTCAAG 422
 QY 361 CGCTGCAAGCGGCGCTGCCAGCTTCCGCAAGTCCGAGCTCCAGCGCGCGCGGCTGCGG 420
 DB 423 CGCTGCAAGCGGCGCTGCCAGCTTCCGCAAGTCCGAGCTCCAGCGCGCGGAGGCTGCGG 482
 QY 421 GACTTCCAGCGCGGAGCGCTTCAAGTCTCGTCAAGTCTGCTTACTTCAAGCGCAATAAT 480
 DB 483 GACTTCCAGCGCGGAGCGCTTCAAGTCTCGTCAAGTCTGCTTACTTCAAGCGCAATAAT 542
 QY 481 CTCGCCAAGCGCTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
 DB 543 CTCGCCAAGCGCTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 602
 QY 541 ATGAGAGGACATGGCGATATTAAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
 DB 603 ATGAGAGGACATGGCGATATTAAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 662

QY 601 CTGGAACCAAGTCATATGAAAGCCCTGTTTCATCCGAGCAGTCGGGCATACAAACGGTGAG 660
 DB 663 CTGGAACCAAGTCATATGAAAGCCCTGTTTCATCCGAGCAGTCGGGCATACAAACGGTGAG 722
 QY 661 AACTGGAGAACATCCATCACAGACATGGAGCTGGCCCTTCCGACCTTCTTCAAGGCTTT 720
 DB 723 AACTGGAGAACATCCATCACAGACATGGAGCTGGCCCTTCCGACCTTCTTCAAGGCTTT 782
 QY 721 TAGCAGTGTCTCGCAGCCTCGAGGGTTCAGGGGATCAAGGACTTCAAGGATTTCTAC 780
 DB 783 TAGCAGTGTCTCGCAGCCTCGAGGGTTCAGGGGATCAAGGACTTCAAGGATTTCTAC 842
 QY 781 CTTTCATACAGACATCATTTATGATAGTCTTGGATGCAAAATACAGTGTGAAGAGAC 840
 DB 843 CTTTCATACAGACATCATTTATGATAGTCTTGGATGCAAAATACAGTGTGAAGAGAC 902
 QY 841 CTCACCCCAAGTATAGGAGGCTATCCGGTTGAGAAATTTGGCTACCATGATGATCATTTAC 900
 DB 903 CTCACCCCAAGTATAGGAGGCTATCCGGTTGAGAAATTTGGCTACCATGATGATCATTTAC 962
 QY 901 TTGCAGTTTGCCTATTATTAAGTTGAACGCTTGAAGATGACGCCCTTGTGCGCTCAGC 960
 DB 963 TTGCAGTTTGCCTATTATTAAGTTGAACGCTTGAAGATGACGCCCTTGTGCGCTCAGC 1022
 QY 961 TATCTGCTCTTTGATCAGAAATGCAAGGTCATGCGAGCAGAACTGCTGTATTACAGTAC 1020
 DB 1023 TATCTGCTCTTTGATCAGAAATGCAAGGTCATGCGAGCAGAACTGCTGTATTACAGTAC 1082
 QY 1021 CACAGGACACTTGGGCGCTCTCGGATGAGCACTTCCAGCCAGACCTTGAAGCAGTTCAG 1080
 DB 1083 CACAGGACACTTGGGCGCTCTCGGATGAGCACTTCCAGCCAGACCTTGAAGCAGTTCAG 1142
 QY 1081 TTCTTTAATGTGACACACACTCCAGAGGAGCTGATGACTTTTCTTAAGGAAATATAATG 1140
 DB 1143 TTCTTTAATGTGACACACACTCCAGAGGAGCTGATGACTTTTCTTAAGGAAATATAATG 1202
 QY 1141 GATGATGATGAGGAGAGTCTTGGAAATATGTTGGATGACCTCTTGGAACTGAGGAGACC 1200
 DB 1203 GATGATGATGAGGAGAGTCTTGGAAATATGTTGGATGACCTCTTGGAACTGAGGAGACC 1262
 QY 1201 AGC 1203
 DB 1263 AGC 1265

RESULT 3
 BC008745
 LOCUS
 DEFINITION Homo sapiens, cartilage associated protein, clone MGC:1926
 IMAGE:3347384, mRNA, complete cds.
 ACCESSION BC008745
 VERSION BC008745.1 GI:14250580
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1439)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (25-MAY-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-3590,
 USA
 REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 COMMENT Contact: MGC help desk
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: DCTD/DTF
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
 cDNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov>
<http://nisc.mcgenetri.nih.gov>
 Shchavchenko, Y., Wachsberg, K.D., Weststrom-Sternberg, S.M.,
 Benjamin, B., Blackley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
 Lim, M., Maduro, Q.L., Masiello, C., Mastrian, S.D., McCluskey, J.C.,
 McDowell, J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J.,
 Tionson, E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,
 Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: ISAL Plate: 5 Row: i Column: 24
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA qi: 3687321.

FEATURES
SOURCE

```

1. .1439
location/sequences
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="LocusID:10491"
/db_xref="taxon:9606"
/clone="MGC:1936 IMAGE:3347384"
/xisue_type="kidney, renal cell adenocarcinoma"
/clone_lib="NH MGC 14"
/lab_host="DH10B-R"
/note="Vector: pOT7"
36. .1241

```

CDS

```

362 a      415 C      375 G      287 t
BASE COUNT
ORIGIN
0: 14241
/codon_start=1
/product="cartilage associated protein"
/protein_id="AAH08745.1"
/db_xref="GI:14250581"
/translations="MEFGRRGAALLLALLCVACLRAGRAQYERYSRFSRPRDELMPLE
EAGYRHALDKYGEHWAESVGYLEISLRHLRLLRDSEAFCHNCSSAQPPEAPAGLAS
YELRFLGGLLRHAHCLPKQSCOGLPAPROSPORDYDLAQFREPVPYQFAPYKAVN
LPKAIAAATFTLKKHDDENMKRNMAYYKSLPGAEYIIDLKTKSYESLFRKAYRN
GEMRTSYIDMEALDLDPFKAFYCEACGEGREIKDFKDFYLSADH7YVEVLEKIQ
CEENLPV7GGYVEVFATVHYLQFAAYKLNDAKAPCAVSYLLFPQNDKQWQON
LVYQYRHTWGLSDSEHFQRPPEAVQFVNVTTLKELYDFAKENIMDDDEGEVVEYVD
DLLELEETS"

```

Query Match 99.5%; Score 1197.4; DB 9; Length 1439;
Best Local Similarity 99.6%; Pred. No. 1.6e-181;
Matches 1198; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

1	Qy	ATGAGCCGGGGCGCCGGGGGGCCGGGGGCTCTACTGCGTGTGTGCGTGGGCTCGCG	60
36	Db	ATGAGCCGGGGCGCCGGGGGGCGGGGCTCTAGCGTGTGTGTGCGTGGGCTCGCG	95
61	Qy	CTGCGGCGGGCGCGCCCAATACGACGCTACAGCTTCGGAGCTTCCACGGGACGAG	120
96	Db	CTGCGGCGGGCGCGCCCAATACGAACTACAGCTTCGGAGCTTCCACGGGACGAG	155
121	Qy	CTGATCCGCTCAGTTCGGGCTTACCGGCGACGGCGCTGGACAAGTACAGCGCGGAGCACTGG	180
156	Db	CTGATCCGCTCAGTTCGGGCTTACCGGCGACGGCGCTGGACAAGTACAGCGCGGAGCACTGG	215
181	Qy	GCGAGAGGCTGGGCTACCTGGAGATAGCGCTCGGCTGCAACGCTTGTGGCGGACAGC	240
216	Db	GCGAGAGGCTGGGCTACCTGGAGATCAGCGCTCGGCTGCAACGCTTGTGGCGGACAGC	275
241	Qy	GAGGCGCTTCGCCACCGCAACTCAGCGCGCGCGCGACGCCGAGCGCGCGCGCGCTC	300
276	Db	GAGGCGCTTCGCCACCGCAACTCAGCGCGCGCGCGCGACGCCGAGCGCGCGCGCGCTC	335
301	Qy	GCCAGCTATCCGAGCTGGCGCTTCTGGGGGGCGCTGTGCGCGCGCGCACTGCCTCAAAG	360
336	Db	GCCAGCTATCCGAGCTGGCGCTTCTGGGGGGCGCTGTGCGCGCGCGCACTGCCTCAG	395
361	Qy	CGGTGCAAGAGGGCTGCAGGCTTCGCGAGTCCCGACGCCAGCGCGGAGTGTGGCG	420
396	Db	CGGTGCAAGAGGGCTGCAGGCTTCGCGAGTCCCGACGCCAGCGCGGAGTGTGTGGCG	455

[illegible]

ORIGIN

Query Match 98.7%; Score 1187; DB 9; Length 2010;
Best Local Similarity 99.6%; Pred. No. 7e-180;
Matches 1198; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

```
QY 1 ATGAGCGCGGGCGCGCGGGGCGCGCGCTGCTAGCGCTGCTGCGTGGCGTGGCG 60
Dd |||
QY 96 ATGAGCGCGGGCGCGCGGGGCGCGCGCTGCTAGCGCTGCTGCGTGGCGTGGCG 155
Dd |||
QY 61 CTGCGCGCGGGCGCGCGCGCGCGCGCTGCTAGCGCTGCTGCGTGGCGTGGCG 120
Dd |||
QY 156 CTGCGCGCGGGCGCGCGCGCGCGCGCTGCTAGCGCTGCTGCGTGGCGTGGCG 215
QY 121 CTGATGCGCGCTGAGTGGCGCTTACCGGCGCGCTGCGCGCTGCGCGCTGCG 180
Dd |||
QY 216 CTGATGCGCGCTGAGTGGCGCTTACCGGCGCGCTGCGCGCTGCGCGCTGCG 275
QY 181 GCCGAGAGCGTGGCTGCTGAGATGAGCTGCGGCTGCGCGCTGCGCGCTGCG 240
Dd |||
QY 276 GCCGAGAGCGTGGCTGCTGAGATGAGCTGCGGCTGCGCGCTGCGCGCTGCG 335
QY 241 GAGGCTTCTGCGACCGCACTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCTC 300
Dd |||
QY 336 GAGGCTTCTGCGACCGCACTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCTC 395
QY 301 GCCAGTATCCGAGCTGCGCTTCTGCGGCGCGCTGCGCGCGCGCGCGCGCTC 360
Dd |||
QY 396 GCCAGTATCCGAGCTGCGCTTCTGCGGCGCGCTGCGCGCGCGCGCGCGCTC 455
QY 361 CGTGTGAAGCGGCGCTGCGAGCTTCCGCGCGCTGCGCGCGCGCGCGCGCTC 420
Dd |||
QY 456 CGTGTGAAGCGGCGCTGCGAGCTTCCGCGCGCTGCGCGCGCGCGCGCGCTC 515
QY 421 GACTTCCAGCGCGCGCGCGCTGCGAGCTTCCGCGCGCTGCGCGCGCGCGCGCTC 480
Dd |||
QY 516 GACTTCCAGCGCGCGCGCGCTGCGAGCTTCCGCGCGCTGCGCGCGCGCGCGCTC 575
QY 481 TCCCGCAAGCGCGCTGCGAGCTTCCGCGCGCTGCGCGCGCTGCGCGCGCTC 540
Dd |||
QY 576 TCCCGCAAGCGCGCTGCGAGCTTCCGCGCGCTGCGCGCGCTGCGCGCGCTC 635
QY 541 ATGAGAGGAGCACTGGCATATTAAGAGCGCTGCGCGCGCTGCGCGCGCTC 600
Dd |||
QY 636 ATGAGAGGAGCACTGGCATATTAAGAGCGCTGCGCGCGCTGCGCGCGCTC 695
QY 601 CTGGAACCAAGTATGAGAGCTTCTGAGAGCTTCTGAGAGCTTCTGAGAGCT 660
Dd |||
QY 696 CTGGAACCAAGTATGAGAGCTTCTGAGAGCTTCTGAGAGCTTCTGAGAGCT 754
QY 661 AACTGGAGCAATCCATCAGAGATGAGAGCTTCTGAGAGCTTCTGAGAGCT 720
Dd |||
QY 755 AACTGGAGCAATCCATCAGAGATGAGAGCTTCTGAGAGCTTCTGAGAGCT 814
QY 721 TACGAGTGTCTGCGAGCTGCGAGCTTCCAGGAGATCAGGAGCTTCCAGGAG 780
Dd |||
QY 815 TACGAGTGTCTGCGAGCTTCCAGGAGATCAGGAGCTTCCAGGAGCTTCCAGG 874
QY 781 CTTTCCATGAGAGATCATTATGAGAGCTTCCAGGAGATCAGGAGCTTCCAGG 840
Dd |||
QY 875 CTTTCCATGAGAGATCATTATGAGAGCTTCCAGGAGATCAGGAGCTTCCAGG 934
QY 841 CTACCCCGAGTTATGAGAGCTTCCAGGAGATCAGGAGCTTCCAGGAGCTTCCAG 900
Dd |||
QY 935 CTACCCCGAGTTATGAGAGCTTCCAGGAGATCAGGAGCTTCCAGGAGCTTCCAG 994
QY 901 TTGAGTTGCTGCTATTAAGTGAAGCTTCCAGGAGATCAGGAGCTTCCAGGAG 960
Dd |||
QY 995 TTGAGTTGCTGCTATTAAGTGAAGCTTCCAGGAGATCAGGAGCTTCCAGGAG 1054
QY 961 TATGCTCTTTGATCAGAGTATGAGAGCTTCCAGGAGATCAGGAGCTTCCAGGAG 1020
Dd |||
QY 1055 TATGCTCTTTGATCAGAGTATGAGAGCTTCCAGGAGATCAGGAGCTTCCAGGAG 1114
```

```
QY 1021 CACAGGACACCTTTGGGCGCTCTCGGATGAGCACTTCCAGCCAGACCTGAAGCAGTTCA 1080
Dd |||
QY 1115 CACAGGACACCTTTGGGCGCTCTCGGATGAGCACTTCCAGCCAGACCTGAAGCAGTTCA 1174
Dd |||
QY 1081 TTCTTTAATGTGACCACTCCAGAGGAGCTGTATGACTTTGCTTAAGAAATATAATG 1140
Dd |||
QY 1175 TTCTTTAATGTGACCACTCCAGAGGAGCTGTATGACTTTGCTTAAGAAATATAATG 1234
QY 1141 GATGATGATGAGGAGAAAGTTTGGGAATATGAGGATGACCTTTGGAACCTGGAGGAGCC 1200
Dd |||
QY 1235 GATGATGATGAGGAGAAAGTTTGGGAATATGAGGATGACCTTTGGAACCTGGAGGAGCC 1294
QY 1201 AGC 1203
Dd |||
QY 1295 AGC 1297
```

RESULT 7

AK091772
LOCUS
DEFINITION
Homo sapiens cDNA FLJ34453 fis, clone HLUNG2002429, highly similar
to Homo sapiens mRNA for cartilage-associated protein.
AK091772
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

1922 bp mRNA linear PRI 15-JUL-2002
1
AK091772.1 GI:21750221
Oligo capping, fis (full insert sequence).
Homo sapiens
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Tashiro, H., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S.,
Fukuzumi, Y., Fujimori, Y., Komiyama, M., Sugiyama, T., Irie, R.,
Otsuki, F., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y.,
Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H.,
Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,
Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A.,
Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,
Nagahari, K., Masuho, Y., Nagai, K. and Isogai, T.
NEDO human cDNA sequencing project

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished
2 (bases 1 to 1922)
Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FUJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.

FEATURES

source

1..1922
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HLUNG2002429"
/tissue_type="lung"
/clone_lib="HLUNG2"
/note="Cloning vector: pME18SPL3"
40..1206
/note="unnamed protein product"
/protein_id="BAC03743.1"
/db_xref="GI:21750222"

CDS

/codon_start=1
/protein_id="BAC03743.1"
/db_xref="GI:21750222"
/translation="MEPGRGAAALLALLCVACALRAGRAQYERYSFPRDLMP
ESAYRHALDKYSGEHWAEVSDFCHRCNSAAPOPEAGLASYPELRFLGGLRR
AHLCKRCKGQAPFQSQPSREVLAQFQREPKFLOFAVFKANLPEKATAAHTFL
KHPDDEMKRWAYKSLPGREDYIKLETKSYESLFIKRAVNGENWRTSIDMEL

ALPDPFFAFVECLAACGSEIRIDFKDFYLSIADHYVEVLECKIQEENLTPVIGVP
VERFATMYHLOFAYVKNLKNAPCAVSYLLFDQNDQMOMNLVYVYHRDITWGL
SDEHQPRPEAVFNTVITLQKLSJYFPAKENIMDDGEVEVYVDLLELETS"
413 a 566 c 467 g 476 t
BASE COUNT
ORIGIN

Query Match	92.4%;	Score 1111.4;	DB 9;	Length 1922;
Best Local Similarity	96.5%;	Pred. No. 8.2e-168;		
Matches 1161; Conservative	1;	Mismatches 2;	Indels 39;	Gaps 1;

QY	1	ATGGAGCGGGGCGCGGGGGCGCGCGCTGCTAGCGTGTCTGTGGCTGGCGCTGCGG	60
Db	40	ATGGAGCGGGGCGCGGGGGCGCGCGCTGCTAGCGTGTCTGTGGCTGGCGCTGCGG	99
QY	61	CTGGCGCGGGGCGCGCCCAATACGAACGCTACAGCTTCGGCAGCTTCCACGGGACGAG	120
Db	100	CTGGCGCGGGGCGCGCCCAATACGAACGCTACAGCTTCGGCAGCTTCCACGGGACGAG	159
QY	121	CTGATGCGGCTGAGTCTGGCTTACCGGACACGCGTGGACAAGTACAGCGCGCAGCACTGG	180
Db	160	CTGATGCGGCTGAGTCTGGCTTACCGGACACGCGTGGACAAGTACAGCGCGCAGCACTGG	219
QY	181	GCCGAGAGGCTKXGGCTACTGAGGATCAGCGCTGGGCTGCACCGCTTGCTGGCGCAGACG	240
Db	220	GCCGAGAGGCTKXGGCTACTGAGGATCAGCGCTGGGCTGCACCGCTTGCTGGCGCAGACG	240
QY	241	GAGGCTTCTGCCACCGCAACTGACAGCGCGCGCGCAGCGCGCGCGCGCGCGCTTC	300
Db	241	GAGGCTTCTGCCACCGCAACTGACAGCGCGCGCGCGCAGCGCGCGCGCGCGCTTC	300
QY	301	GCAGCTATCCGAGCTGGGCTCTTTGGGGGCTGCTGGCGCGCGCGCAGCTGCCTCAAG	360
Db	301	GCAGCTATCCGAGCTGGGCTCTTTGGGGGCTGCTGGCGCGCGCGCAGCTGCCTCAAG	360
QY	361	CGGTGAACAGCGGCTGCGAGCTTCGCGCAGTCCGAGCCCGCGCGCGAGGTGCTGGCG	420
Db	361	CGGTGAACAGCGGCTGCGAGCTTCGCGCAGTCCGAGCCCGCGCGCGAGGTGCTGGCG	420
QY	421	GACTTCCAGCGCGGAGCGCTTACAAGTTCTGCAGTTCGCTTACTTCAAGGCAATATAT	480
Db	421	GACTTCCAGCGCGGAGCGCTTACAAGTTCTGCAGTTCGCTTACTTCAAGGCAATATAT	480
QY	481	CTCCCCAAGCCATCGCGCTGCTACACCTTCTACTGAAGCATCCTGATGACGAATG	540
Db	481	CTCCCCAAGCCATCGCGCTGCTACACCTTCTACTGAAGCATCCTGATGACGAATG	540
QY	541	ATGAAGAGGAACATGCGATATTATAGAGCTGCGCTGGTCCGAGACTACATTAAAGNC	600
Db	541	ATGAAGAGGAACATGCGATATTATAGAGCTGCGCTGGTCCGAGACTACATTAAAGAC	600
QY	601	CTGGAAACCAAGTCTATATGAAGCGCTGTTCATCCGAGCAGTGGGGCATACAACTGAG	660
Db	601	CTGGAAACCAAGTCTATATGAAGCGCTGTTCATCCGAGCAGTGGGGCATACAACTGAG	660
QY	661	AACTGGAGAACATCCATCACAGACATGGAGCTGCGCTTCCGAGCTTCTTCAAGCGCTTT	720
Db	661	AACTGGAGAACATCCATCACAGACATGGAGCTGCGCTTCCGAGCTTCTTCAAGCGCTTT	720
QY	721	TACGAGTGTCTCGAGCCTGGGAGGTTCCAGGAGATCAAGACTTCAAGATTCTCTAC	780
Db	721	TACGAGTGTCTCGAGCCTGGGAGGTTCCAGGAGATCAAGACTTCAAGATTCTCTAC	780
QY	781	CTTTCATAGAGATCATTATGTAGAGTTCTGGAATGCAAAATACAGTGTGAAGAGAAC	840
Db	781	CTTTCATAGAGATCATTATGTAGAGTTCTGGAATGCAAAATACAGTGTGAAGAGAAC	840
QY	841	CTCACCCCGATTATAGGAGGCTATCCGGTTGAGAAATTTGTGGCTACCATGTATCATTAC	900
Db	841	CTCACCCCGATTATAGGAGGCTATCCGGTTGAGAAATTTGTGGCTACCATGTATCATTAC	900
QY	901	TTGCAGTTTGCCATTATTAAGTTGAACGCTGAAGAAATGCAGCCCGCTGTGCAGTCAAG	960
Db	901	TTGCAGTTTGCCATTATTAAGTTGAACGCTGAAGAAATGCAGCCCGCTGTGCAGTCAAG	960

Qy	961	TATCTGCTCTTTGATCAGATGACAAAGGTGATCGACGAGAACTCGTGTTATTACCACTAC	1020
Db	961	TATCTGCTCTTTGATCAGATGACAAAGGTGATCGACGAGAACTCGTGTTATTACCACTAC	1020
Qy	1021	CACAGGACACTTGGGGCCTCTCGATGAGCACTTCAGCCCAAGACCTGAAAGCAGTTTCAG	1080
Db	1021	CACAGGACACTTGGGGCCTCTCAGATGAGCACTTCAGCCCAAGACCTGAAAGCAGTTTCAG	1080
Qy	1081	TTCTTTAATGTGACCACTCCAGAAAGGAGCTGTATGACTTTGCTAAGGAAAAATAAATG	1140
Db	1081	TTCTTTAATGTGACCACTCCAGAAAGGAGCTGTATGACTTTGCTAAGGAAAAATAAATG	1140
Qy	1141	GATGATGATCAGGAGGAAGTTGTGGAAATATGTGGATGACCTCTTGGAACTGGGAGGAGACC	1200
Db	1141	GATGATGATCAGGAGGAAGTTGTGGAAATATGTGGATGACCTCTTGGAACTGGGAGGAGACC	1200
Qy	1201	AGC 1203	
Db	1201	AGC 1203	

RESULT 8	
AK092990	AK092990
LOCUS	1903 bp mRNA linear
DEFINITION	Homo sapiens cDNA FLJ35671 fis, clone SPLEN201818.0, highly similar to Homo sapiens mRNA for cartilage-associated protein.
ACCESSION	AK092990
VERSION	AK092990.1 GI:21751708
KEYWORDS	oligo capping; fis (full insert sequence).
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

REFERENCE
AUTHORS
Tashiro, H., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Koniyama, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Ota, T., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai, Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Nagatsuna, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Sugano, S., Nagahari, K., Masuho, Y., Nagai, K. and Isogai, T.
TITLE
NEDO human cDNA sequencing project

REFERENCE	2 (bases 1 to 1903)
AUTHORS	Isogai, T. and Yamamoto, J.
TITLE	Direct Submission
JOURNAL	Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
COMMENT	Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel:81-438-53-3975, Fax:81-438-53-3986) NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

```

BASE COUNT      421 a      553 c      460 g      469 t
/notes=cloning vector: pME18SFL3"
/clone_lib="SPLEN2"
/issue_type="splice"
/clone="SPLEN2018180"
/db_xref="taxon:9606"
/mcl_type="mRNA"
/organism="Homo. sapiens"
source
localdir/Qualifiers
1..1903

```

Query Match	89.5%	Score 1076.6;	DB 9;	Length 1903;
Best Local Similarity	95.4%	Pred. No. 2.9e-162;		

13	CGCCGGGGCCCGCGCGCTGCTAGCGCTGCTGTGCTGCGCTGCGCGCGCGG	72
60	CGCAGCCCTACGCGCGCGTTGCTGCTGCTGCTGCTGCTGCGCGCGCGCGG	119
73	CGCGCCCAATACAAACGCTACAGCTTCCGAGCTTCCACGGAGACAGCTGATCGCGCTC	132
120	CGGGCAGTACAGCGCTACAGCTTCCGAGCTTCCCGGGACAGCTGATGCCGCTC	179
133	GAGTCGGGCTTACCGGCACGCGCTTGGCAAGTACAGCGGCGAGCACCTGGGCGGAGAGCSTK	192
180	GAGTCGGGCTTACCGCCACGCGCTCGACCACTACAGCGGCGAGCACTGGGCGGAGAGCGTG	239
193	GGCTACTGGAGATACGCTCGGGCTGCACCGCTTGTGTGCGACAGCGAGCGCTTCTGC	252
240	GGCTACTGGAGTGAGCCTCGGGCTGCACCGCTTGTGTGCGACAGCGAGGCGCTTCTGC	299
253	CACCGCACTGCAGCGCCGCGCGCAGCCGAGCCGCGCGCGCTCCGCAAGCTATCCC	312
300	CACCGCACTGCAGCGCGCCACGCCAGCACCGCGCCCGCGCGCGCGCGCCACGCCACGCC	359
313	GAGCTCGGCTCTTTGGGGGGCTGTGTGCGCGCGCACTGCTCAAGCGCTGCAGACAG	372
360	GAATACGCGCTCTTTCGGCAGCGTGTGTGCGCGCGCAGTGCTCAAGCGCTGCAGCAG	419
373	GGCTTGCACAGCTTCCGCGAGTCCAGCCACCGCGAGGTGCTGCGGACTTCCAGCGC	432
420	GGCTTGCAGCGCTTCCGCGAGTCCAGCGCGCAGCGCTCAGTGTGCGCGACTTTACGAC	479

[illegible]

[illegible]

RESULT 13			
GGCARPR	GGCARPR	1665 bp	linear
LOCUS	LOCUS		
DEFINITION	G.gallus mRNA for cartilage associated protein.		
ACCESSION	X97607		
VERSION	X97607.1		
	GI:1296525		
			VRT 26-JUN-1997

QY 431 GCCGCGAGCCTACAGTTCTCGAGTTCCTTACTTCAAGCAATATCTCCCAAG 490
Db 443 GCCGAGAGCCCTCAAGTACTCGAGTCCCTACTTCAAGGTAATATCCCTTCCAAAG 502
QY 491 CCATCGCGCTGTCTCACACCTTTCTACTAAGCATCTCTGATGACGAAATGATGAAGGA 550
Db 503 CTATTGCGAGAGCTCACACATTTCTTCTGAAATCATCCAGATGATGAATGATGAAGGA 562
QY 551 ACATGGCATATTATAGAGCCTGCTGTGTCGCCGAGGACTACATTAAGACCTTGGAAACCA 610
Db 563 ATATGGCCTACTATAAGAGCATACCTGATGCTGAGGAGCACATTAAGACTTGGAGACAA 622
QY 611 AGTCATATGAAGCCTTTCATCGAGAGCTGCGGCATACACAGGTGAGAACTGGAGAA 670
Db 623 AGCCTTACGAGATCTTTGTGAGGCTGTGAGACATACATGTGTGACACTGGAGAA 682
QY 671 CATCCATCAGACATGAGAGCTGGCCCTCCGAGCTTCTCAAGCCTTTTACGAGTGTG 730
Db 683 CATCCATCTCAGATATGGAATGGCTCTTCTGATTTCTCAAAACCTACGATGACTGTA 742
QY 731 TCGCAGCTCGAGGTTCCAGGAGATCAAGGACTCAAGATTTCTACCTTCCATAG 790
Db 743 TAGAGCCTGTGAGGCTCCGAGAGATTAAGATTTTAAGACTTCTATCTCTCTATG 802
QY 791 CAGATCATTATGATGAGTCTTGGAAATGCAAAATACAGTGTGA 833
Db 803 CAGATCATTATATTGAATGCTTGCATGCAAAAGTCAGTTCGA 845

RESULT 14

AC112211
LOCUS 182509 bp DNA linear PRI 24-MAY-2002
DEFINITION Homo sapiens chromosome 3 clone RP11-627J17, complete sequence.
ACCESSION AC112211
VERSION AC112211.2 GI:21166208

KEYWORDS
SOURCE
ORGANISM

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 182509)
AUTHORS Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
Saenphimmachak,C., Phelps,K.A., Buckley,D., Raymond,C. and
Haugen,E.D.

TITLE

Direct Submission
Unpublished
2 (bases 1 to 182509)
Kaul,R.K., Olson,M.V., Raymond,C. and Haugen,E.D.
Direct Submission
Submitted (20-FEB-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA

REFERENCE

1 (bases 1 to 182509)
Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
Saenphimmachak,C., Phelps,K.A., Buckley,D., Raymond,C. and
Haugen,E.D.

TITLE

Direct Submission
Submitted (24-MAY-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA

COMMENT

On May 24, 2002 this sequence version replaced gi:18767508.

Genome Center
Center: University of Washington Genome Center
Center Code: UWGC
Web site: <http://www.genome.washington.edu>
Contact: uwgctgs@u.washington.edu
----- Project Information
Center project name: chr-3
Center clone name: RP11-627J17 (bc0502)
----- Summary Statistics
Sequencing vector: plasmid; 100% of reads
Chemistry: Dye-terminator ET; 50% of reads
Chemistry: Dye-terminator Big Dye; 50% of reads
Assembly program: Phrap; version 0.990319

Consensus quality: 182361 bases at least Q40
Consensus quality: 182497 bases at least Q30
Consensus quality: 182509 bases at least Q20
Insert size: 182509; sum-of-contigs
Quality coverage: 7.9x in Q20 bases; sum-of-contigs

Overlapping Sequences:

5: Mapping in progress
3: Mapping in progress

Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp.

Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:

This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

EcoRI

SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt
8696	8777	4710	4671	2282	2251
6	<800	6382	6514	2067	2068
8311	8264	512	<800	5626	5794
3794	3847	449	<800	4289	4117
517	<800	1417	1461	4382	4412
5654	5717	5366	5126	4594	4580
3530	3552	590	<800	10637	10624
2635	2737	3606	3651	4577	4580
3532	3552	274	<800	246	<800
12069	11650	4570	4539	933	948
8046	8264	9049	8995	15529	15666
2974	3009	4374	4320	184	<800
3451	3552	4978	5126	4296	4283
2090	2126	11396	11164	2736	2818

BglII

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbiology.org>
 contact: anadan@systemsbiology.org
 Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julian Greene, Mark Ketteman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 7 Row: f Column: 22
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 5454037.

FEATURES
source

```

1. .1889
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:10609"
/db_xref="taxon:9608"
/clone="MGC:1322 IMAGE:3510406"
/tissue_type="Brain, neuroblastoma"
/clone_lib="NIH MGC 19"
/lab_host="DH10B-β"
/note="Vector: pOTB7"
47. .1360
/codon_start=1
/product="nucleolar autoantigen (55kD) similar to rat
synaptonemal complex protein"
/protein_id="AAH01047.1"
/db_xref="GI:126514443"
/translation="MARVAGLWLLSGAQYKYKFRGPPEDLMPALAAAYGHAL
EQYEGWRSAYLEALRHLRLDSFAFCHANCSPAPAAKPDGGRADWEACE
LRFGVLEAARLCKRCLTPAFQYVPPOLLRPDQSPFLPYOVLHYALFKANRLK
AAVAAYTFGLQNPGRKLTAKYLVNPGMLDVAESLTLEAQYEAFLRAVKLYNSG
DFRSTDMERALSLEYLAPARCAGACEGAQVDFKDFPAIADLFAESLQCKYDCE
ANLIPNVGGVFVDFKVFATMYTHYLQFPAYYKLVNDVPOAKRASAATYMLDFPKDQSVQWQMLY
YIFRHRARGLSEEDFQFEEMALWYHNQTAERLELFTHYLQSDMELESTETPFL
EPEDLADSAFEGEGYDFEGEAMVQWDMWQDPAKGDGDAEAEPELA"
999 a 570 c 611 g 309 t
COUNT

```

Query Match	38.8%	Score	467;	DB	9;	Length	1889;
Best Local Similarity	64.1%	Prod. No.	S.7e-65;				
Matches	755;	Conservative	2;	Mismatches	402;	Indels	18; Gaps 3;

2y	8	CGGGGGCGGGGGGGCGGGGGCTGCTAGCGCTGCTGCGTGGGCTGCGCGCTGGCGG	67
bb	30	CGGGGGCGGGGGGCGATGCTCGGGTGGCGTGGGGCTGCTGTGTTGCTTCTGGGCA	89
2y	68	CGGGGGCGGGCCCAATACGAACGCTAAGCTTCCGAGCTTCCCAAGGACGAGCTGATGC	127
bb	90	GCGCCGGGGCGAGTACGAAGAAGTACAGCTTCCGGGGCTTCCGCGCGGAGCACTGATGC	149
2y	128	CGTCTGAGTCTACCGCATACCGGCGGGCTGGACAATCAGCGGGGAGACACTGGGCGCAGA	187
bb	150	CGTGGCGGGGGCTACGGGCGACGCTCTGGAGCACTACAGGAGAGAGCTGGGCGCAGA	209
2y	188	GCSTKGGTACCTGGAGATCAGCGCTCGCGCTGCACCGCTTCTCGCGACAGCGAGGCGCT	247
bb	210	GCGGCGGCTACCTGGAGGGGGGGGCTCGGGCTGCACCGCTCTCTCGCGACAGCGAGGCGCT	269
2y	248	TCTGCCACCGCAACTCGAGG-----CGCGCCGCGAGCCCGAGCCCGCCCGCGGCG	298
bb	270	TCTGCCACGCCAACTCGAGCGGGCCCGCGCCCGCGGCGCAAGCCCGATCTCCGACGGGCGCC	329
2y	299	TGCCCACTATC-----CCGAGCTTGGCGCTCTTTCGGGGGGCTCTCTCGCGCGCGCGCACT	352

Search completed: January 30, 2004, 14:49:17
Job time : 4386.75 secs

330	GGCGACAGAGTGGGCTTCGGAGCTTCCTTGGCCGGCGTCTCGAGAGCGAGCCGCT	389
353	GGCTAAGCGCTGCAAGCAGGGCCCTGCCAGCTTCCTCGCCAGTCCCAAGCCAGCCGCGAGG	412
390	GGCTGGGGCGCTGCAGCGGAGCGCTGCCCGCTTCCTCAGGTGCCCTACCCGCGCGCGCAGC	449
413	TGCTGGCGGACTTCAGCGCCCGGAGCCCTACAAGTTCCTCGAGTTCGCTTACTTCAAGG	472
450	TGCTGCGTGACTTCAGAGCGCCCTGCCCTACAGTACCTGCACTACGGCGCTGTTCAAGG	509
473	CAAAATAATCTCCCAAGGCCATGCGCGTGCTCAACCTTTCTACTGAAGCATCTCTGATG	532
510	CTAAACCGGCTGGAGAAGCGGTGGGGGGCGCTACACCTTCTCCAGAGGAACCCGGAAGC	569
533	ACGAAATGATGAAGAGGAACATGGCATATTATAA---GAGCCTGCTGTGCGGAGGACT	589
570	ACGAGCTGACCGCCAAGTATCTCAACTACTATCAGGGGATGCTGAGCGTGCAGCGAGT	629
590	ACATTAAGACCTGGAACCAAGTCAATATGAAGCCTGTTATCCGAGCAGTGCGGGCAT	649
630	CCCTCAGCGACCTAGAGGCCACGCCCTACGAGCGCTGTTCTCCGGGCTGTGAAGCTCT	689
650	ACAAACGGTGAACCTGAGAACATCCATCACAGACATGAGTGTGCCCTTCCGCACTTCT	709
690	ACAAACAGCGGGATTTCCGACGACGACGAGGACATGGAGCGGCGCTTGTCAAGATACC	749
710	TCAAGACCTTTTACAGTGTCTCGAGCGCTCGGAGGTTCCAGGAGATCAAGGACTTCA	769
750	TGGCAGTCTTTGCGCGGTGCTCGCGCGCTGTGAAGGGGCCCATGAGCAGGTGGAATTCA	809
770	AGGATTTCTACTCTTTCCATAGCAGATCATTTATGTAGAAGTTCGGAATGCAAAATACAGT	829
810	AGGACTTCTACCGGSCCATAGCAGATCTTTGCAGAGTCCCTGCAGTCAAGGTGGACT	869
830	GTGAGAGAACTCACCCGACTTAGAGGGTATCCGGTTAGAGAAATTTGTGGCTACCA	889
870	GTGAGCCAAATTTGACCCCAATTTGGTGGCTACTTCGTGACAAGTTCGTGGCCACCA	929
890	TGTATCAATTACTTGCAGTTTCCTATTAATAAGTTGAACGACCTGAAGAATGCAGGCCCT	949
930	TGTACCACTACTGCAGTTTGCCTACTATAAGTTGAATGATGTGGCGCAGGCTGCCGCA	989
950	GTGAGCTCAGCTATCTGCTCTTTGATCAGATGACAGGTATCAGCAGAACCTGGTGT	1009
990	GGCGCGCAGCTCATGCTCTTTGCACCCCAAGGACAGCGTATGAGCAGAACCTGGTGT	1049
1010	ATTACCACTACCAGGGACACTTTGGGSCCTCTCGGATGAGCACTTCAGGCCACGACCTG	1069
1050	ATTACCGGTTCCACCGGCTCGCTGGGSCCTGGAAGAGGAGGACTTCAGCCCCGGGAGG	1109
1070	AAGCAGTTCAGTCTTTTAATGTGACCACTCCAGAGGGAGCTGATGACTTTGCTAAGG	1129
1110	AGGCATGCTCTACCAACAACGACCGCGAGCTGCGGAGCTGCTGAGTTTCACCACA	1169
1130	AAATATAATGATGATGATGAGGAGAAGTTGTGGA	1166
1170	TGTACCTGCAGTCAGATGATGAGATGAGCTGGAGGA	1206

Search completed: January 30, 2004, 14:49:17
Job time : 4386.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2004, 07:10:40 ; Search time 14098.5 Seconds
(without alignments)
11232.506 Million cell updates/sec

Title: US-09-729-674-1
Perfect score: 3871
Sequence: 1 ttctcttccctcccttt.....aataaaaaaaaaaaaaaa 3871

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813396 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vt.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_on.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_scs.*
28: em_un.*
29: em_vt.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3870.2	100.0	3871	6	BD140451 Secreted
2	2619.6	67.7	182509	9	AC112211 Homo sapi
3	2605.2	67.3	166485	9	AC123900 Homo sapi
4	2237	57.8	2307	9	AJ006470 Homo sapi
5	1980.4	50.6	2010	9	AK075401 Homo sapi
6	1955.4	50.5	1961	6	BD156517 Homo sapi
7	1955.4	50.5	1961	6	BD156517 Primer fo
8	1869.4	48.3	1922	9	AK091772 Homo sapi
9	1798.8	46.5	1903	9	AK092990 Homo sapi
10	1656	42.8	1810	9	AK097650 Homo sapi
11	1370.4	35.4	1439	9	EC008745 Homo sapi
12	944	24.4	1685	10	MMAJ6469 Mus muscu
13	940	24.3	1663	10	BC049890 Mus muscu
14	695.6	18.0	818	6	BD148800 Primer fo
15	510.2	13.2	64706	2	AC136358 Homo sapi
16	497.6	12.9	1665	5	GACARPR1 X97607 G. gallus mr
17	467	12.1	1889	9	BC001047 Homo sapi
18	467	12.1	2037	9	BC007942 Homo sapi
19	467	12.1	2443	6	AX714040 Sequence
20	467	12.1	2443	9	AK056085 Homo sapi
21	466.6	12.1	1745	9	BC011701 Homo sapi
22	465.8	12.0	2347	6	AX335490 Sequence
23	465.8	12.0	2347	9	HSU47621 Homo sapien
24	465.4	12.0	2079	9	HSU47621 Homo sapien
25	428.5	11.1	1903	10	BC031856 Homo sapi
26	403	10.4	1407	10	BC031856 Mus muscu
27	392.4	10.1	394	6	AX408806 Sequence
28	369.4	9.5	400	11	G27784 human STS S
29	355.6	9.2	407	6	AX150162 Sequence
30	352.2	9.1	226615	2	AC122652 Rattus no
31	340	8.8	351	6	BD153832 Primer fo
32	313.6	8.1	330	6	AX150140 Sequence
33	283.4	6.9	274	6	AX333823 Sequence
34	258.6	6.7	279	6	AX261032 Sequence
35	252.2	6.5	157326	9	AC008810 Homo sapi
36	250.6	6.5	257967	2	AL365337 Mus muscu
37	249.6	6.4	164217	9	AC093264 Homo sapi
38	249	6.4	200240	2	AC011618 Homo sapi
39	247.8	6.4	157375	9	AC025207 Homo sapi
40	247.8	6.4	178420	9	AP003548 Homo sapi
41	247.8	6.4	199569	9	AC016250 Homo sapi
42	247.8	6.4	211076	2	AC012214 Homo sapi
43	247.4	6.4	189533	9	AC025678 Homo sapi
44	246.6	6.4	115136	9	AL390994 Human DNA
45	245.8	6.3	169089	9	AC008751 Homo sapi

ALIGNMENTS

RESULT 1
LOCUS BD140451
DEFINITION Secreted proteins and polynucleotides encoding them.
ACCESSION BD140451
VERSION BD140451.1 GI:23235396
KEYWORDS JP 2002506611-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3871)
AUTHORS Jacobs,K., McCoy,J.M., Lavallie,E.R., Racie,L.A.C., Evans,C.,
Merberg,D., Treacy,M., Agostino,M.J., II,R.J.S., Wong,G.G.,
Clark,H.F. and Fechtel,K.

1681 DB CAGGCTGCTCGAACTCTTGACCTCAGATGATCACTGCGCTGGCGCTCCCAAGTGC 1740
1741 QY TGGGATTA CAGCGTGAGCCACATGCCGCGCTCTTCTCACCTTTACACCTGCTCTCT 1800
1741 DB TGGGATTA CAGCGTGAGCCACATGCCGCGCTCTTCTCACCTTTACACCTGCTCTCT 1800
1801 QY TATCTCTACATCTGTTTTCACACCTTCATCCCTGCTCTTCTCATGTTTACACCTGCTCT 1860
1801 DB TATCTCTACATCTGTTTTCACACCTTCATCCCTGCTCTTCTCATGTTTACACCTGCTCT 1860
1861 QY CCCATGTTTCATAGCTGCTCTTCTTACCATTTTGGTTTGAAGGCGAGTCTTCTCTGGCTTG 1920
1861 DB CCCATGTTTCATAGCTGCTCTTCTTACCATTTTGGTTTGAAGGCGAGTCTTCTCTGGCTTG 1920
1921 QY TTTTCTTCTTTTCCAGAAATCAGTATATTTTAAATAAGAAATCTTCTAGAA 1980
1921 DB TTTTCTTCTTTTCCAGAAATCAGTATATTTTAAATAAGAAATCTTCTAGAA 1980
1981 QY GATGATAATTTGTAATCTCTTGGCTTATTTGCTTTTCCAGATTTTGTCTCTCTTTC 2040
1981 DB GATGATAATTTGTAATCTCTTGGCTTATTTGCTTTTCCAGATTTTGTCTCTCTTTC 2040
2041 QY TCCCATCCGGAAAGATGTTGGAAGACATAGGCTTAAATTTCTCCAGCTTCAATGCTC 2100
2041 DB TCCCATCCGGAAAGATGTTGGAAGACATAGGCTTAAATTTCTCCAGCTTCAATGCTC 2100
2101 QY TTCACCTTGGTCTGACTTGTACCAATCTAGCACCCACTGAAACACAGTTGAGTAGAG 2160
2101 DB TTCACCTTGGTCTGACTTGTACCAATCTAGCACCCACTGAAACACAGTTGAGTAGAG 2160
2161 QY TGTAGAGTGCAGAAATGTTGGCTTTTGGCCCCACTTTGCACTCTCCAAATTAACAGGTTGG 2220
2161 DB TGTAGAGTGCAGAAATGTTGGCTTTTGGCCCCACTTTGCACTCTCCAAATTAACAGGTTGG 2220
2221 QY CCGATCCCATTTGAGGACAACTGCTAGTTTATAGTCTCCGAGTTTGGAAAGGAAGAAGC 2280
2221 DB CCGATCCCATTTGAGGACAACTGCTAGTTTATAGTCTCCGAGTTTGGAAAGGAAGAAGC 2280
2281 QY CAGAGCTCTAGTTTTCATTCATCTTTCAGTAAATATTTATTTGAGTACCTACTGCTGTC 2340
2281 DB CAGAGCTCTAGTTTTCATTCATCTTTCAGTAAATATTTATTTGAGTACCTACTGCTGTC 2340
2341 QY TAGGATTTGACTGGAACCTAGATATCTTACAGAAATTAACAGGAAAGTTTCCCTGTCT 2400
2341 DB TAGGATTTGACTGGAACCTAGATATCTTACAGAAATTAACAGGAAAGTTTCCCTGTCT 2400
2401 QY CATGAGCTTACATTTACAGGGAAGAGATAGCCATACATAGGAATAAATATATAC 2460
2401 DB CATGAGCTTACATTTACAGGGAAGAGATAGCCATACATAGGAATAAATATATAC 2460
2461 QY AAGGTATCATGTAGTGAATTTGCTGTGGAAATAAAGCAGGAGGAGGAGTAAAGAA 2520
2461 DB AAGGTATCATGTAGTGAATTTGCTGTGGAAATAAAGCAGGAGGAGGAGTAAAGAA 2520
2521 QY TCCTGGAGATGAGCTGAGTTTAAATGGGGCTCTACTGGGAATGTGAGCTGAGAGA 2580
2521 DB TCCTGGAGATGAGCTGAGTTTAAATGGGGCTCTACTGGGAATGTGAGCTGAGAGA 2580
2581 QY GACGTTAGGGAAGTGGATCTTGGCAAGGCAATCCAGGAGAGGAAACAAGATGTGCACTG 2640
2581 DB GACGTTAGGGAAGTGGATCTTGGCAAGGCAATCCAGGAGAGGAAACAAGATGTGCACTG 2640
2641 QY CCCAAAGTGAAGACTTGTCTCTAGTGTGAGGAAAGAGCAGGAGACCAAGCAGAGTGC 2700
2641 DB CCCAAAGTGAAGACTTGTCTCTAGTGTGAGGAAAGAGCAGGAGACCAAGCAGAGTGC 2700
2701 QY TGGGAGGGGTAGAAATGGAAGGAGCGGCTGGGAGGACAGTGTGTGGAGGCGCTTG 2760
2701 DB TGGGAGGGGTAGAAATGGAAGGAGCGGCTGGGAGGACAGTGTGTGGAGGCGCTTG 2760
2761 QY CTTCTGCTAGTGAAGTGGGAACCACTGGAGGGTTTGAACAGAGGAGTGCCTTGAATGAT 2820
2761 DB CTTCTGCTAGTGAAGTGGGAACCACTGGAGGGTTTGAACAGAGGAGTGCCTTGAATGAT 2820

2821 QY TTATATTTTCAAGGGTCATTTAGCTGCAATATTTGTAATACTTTTAGTGGAAGGGC 2880
2821 DB TTATATTTTCAAGGGTCATTTAGCTGCAATATTTGTAATACTTTTAGTGGAAGGGC 2880
2881 QY AGAAGGAAGAGGAAGAACCTGTTTAGAAGCTACTGCAAGTTTCCAGGCTTGGCCCTGGC 2940
2881 DB AGAAGGAAGAGGAAGAACCTGTTTAGAAGCTACTGCAAGTTTCCAGGCTTGGCCCTGGC 2940
2941 QY CACAGCAACAGCAGTGTCAATATCTAGATTTTATTTTGAAGAGCCCAATAGGATTTGC 3000
2941 DB CACAGCAACAGCAGTGTCAATATCTAGATTTTATTTTGAAGAGCCCAATAGGATTTGC 3000
3001 QY TGAGAGTTTGAATGTGAGTGTAAAGAGGAAGAGTAAATGATGACATTTAAGGTTTTC 3060
3001 DB TGAGAGTTTGAATGTGAGTGTAAAGAGGAAGAGTAAATGATGACATTTAAGGTTTTC 3060
3061 QY GCCTGAATAGCAGGAAGATGGAGTTTACCACTTACTGAAATAGGAGAGTGGCTGGGT 3120
3061 DB GCCTGAATAGCAGGAAGATGGAGTTTACCACTTACTGAAATAGGAGAGTGGCTGGGT 3120
3121 QY AAGTAAAGGAATTTGGTGCAGAGCAGGCTCTCTGTGTTGGAATGGAGGTTCTGGCTGCA 3180
3121 DB AAGTAAAGGAATTTGGTGCAGAGCAGGCTCTCTGTGTTGGAATGGAGGTTCTGGCTGCA 3180
3181 QY AATCAAAGTGGAGATTTCTCTCAGGTCAGGCTGCGAGCAGAGCTCGAGACAGGATCTGAA 3240
3181 DB AATCAAAGTGGAGATTTCTCTCAGGTCAGGCTGCGAGCAGAGCTCGAGACAGGATCTGAA 3240
3241 QY TGCACCTTGGTTTATTTGTTGGGGGCTCTCAGAGGAACCTGTGAAAGCCTTTATCAGTC 3300
3241 DB TGCACCTTGGTTTATTTGTTGGGGGCTCTCAGAGGAACCTGTGAAAGCCTTTATCAGTC 3300
3301 QY ATTTATTTGCTGTGAGAAATTTCTCTGGAGTGTGGGTACATTTGAAGCAAGTACTTCA 3360
3301 DB ATTTATTTGCTGTGAGAAATTTCTCTGGAGTGTGGGTACATTTGAAGCAAGTACTTCA 3360
3361 QY GTTGGGGCAAGTCTCTGAAAGAGGCTGTAGGAGCTCTGGCAGCTACCATCGTGGTAG 3420
3361 DB GTTGGGGCAAGTCTCTGAAAGAGGCTGTAGGAGCTCTGGCAGCTACCATCGTGGTAG 3420
3421 QY TGTGTTGGGGTGGGGTCTCTGGCACTGCTGTGTAAGGATCTGGCAGGCAACCA 3480
3421 DB TGTGTTGGGGTGGGGTCTCTGGCACTGCTGTGTAAGGATCTGGCAGGCAACCA 3480
3481 QY GCGCCCCCTACTGAACCATCAGCATGTGCTGGCAATTTAAAGCCATGCACTGGAAGGGC 3540
3481 DB GCGCCCCCTACTGAACCATCAGCATGTGCTGGCAATTTAAAGCCATGCACTGGAAGGGC 3540
3541 QY CACTGAGATTTGCTCTGAGTATTTACTGAGAACCAAGAGAGCCATGGATGGAGCCC 3600
3541 DB CACTGAGATTTGCTCTGAGTATTTACTGAGAACCAAGAGAGCCATGGATGGAGCCC 3600
3601 QY TTGGGCTCTCTGGGAAATGGGAAATCAGCCAAAGGACTGAGAGGAGTTACCTTAAGGTC 3660
3601 DB TTGGGCTCTCTGGGAAATGGGAAATCAGCCAAAGGACTGAGAGGAGTTACCTTAAGGTC 3660
3661 QY AGAGAAAACCAAGAGAGTGTGTGTTCTGGAAGCTGAGCTTTCTTTTCAACCTCATTC 3720
3661 DB AGAGAAAACCAAGAGAGTGTGTGTTCTGGAAGCTGAGCTTTCTTTTCAACCTCATTC 3720
3721 QY CTTCTCCAAATAAGCCACTTTGTGTAGTTGGGCCCCCTCCAGGGTTGAAGCAAGAGGAGA 3780
3721 DB CTTCTCCAAATAAGCCACTTTGTGTAGTTGGGCCCCCTCCAGGGTTGAAGCAAGAGGAGA 3780
3781 QY AAGGCACAGCGTTTGGGAAACAGACTTTTCTGCAATAGCTTGGGAAGGATTAAGAGGA 3840
3781 DB AAGGCACAGCGTTTGGGAAACAGACTTTTCTGCAATAGCTTGGGAAGGATTAAGAGGA 3840
3841 QY TAGAGTGTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 3871
3841 DB TAGAGTGTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 3871

RESULT 2
 AC112211
 LOCUS
 DEFINITION Homo sapiens chromosome 3 clone RP11-627J17, complete sequence.
 AC112211
 ACCESSION
 VERSION AC112211.2 GI:21166208
 KEYWORDS HTG.
 SOURCE
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 182509)
 Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z., Saenphimmachak, C., Phelps, K.A., Buckley, D., Raymond, C. and Haugen, E.D.
 Direct Submission
 Unpublished
 2 (bases 1 to 182509)
 Kaul, R.K., Olson, M.V., Raymond, C. and Haugen, E.D.
 Direct Submission
 Submitted (20-FEB-2002) Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA
 3 (bases 1 to 182509)
 Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z., Saenphimmachak, C., Phelps, K.A., Buckley, D., Raymond, C. and Haugen, E.D.
 Direct Submission
 Submitted (24-MAY-2002) Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA
 On May 24, 2002 this sequence version replaced gi:18767508.

 Center: University of Washington Genome Center
 Center Code: UWGC
 Web site: <http://www.genome.washington.edu>
 Contact: uwgctgs@u.washington.edu

 Center project name: chr-3
 Center clone name: RP11-627J17 (bc0502)

 Summary Statistics
 Sequencing vector: plasmid; 100% of reads
 Chemistry: Dye-terminator ET; 50% of reads
 Assembly: Dye-terminator Big Dye; 50% of reads
 Consensus quality: Phrap; version 0.990319
 Consensus quality: 182361 bases at least Q40
 Consensus quality: 182497 bases at least Q30
 Consensus quality: 182509 bases at least Q20
 Insert size: 182509; sum-of-contigs
 Quality coverage: 7.9x in Q20 bases; sum-of-contigs

 Overlapping Sequences:
 5': Mapping in progress
 3': Mapping in progress

 Sequence Quality Assessment:
 This entry has been annotated with sequence quality estimates computed by the Phrap assembly program.
 All manually edited bases have been reduced to quality zero.
 Quality levels above 40 are expected to have less than 1 error in 10,000 bp.
 Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

 This sequence was finished as follows unless otherwise noted:
 all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

 Sequence Validation:

This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

ECORI	HindIII	BglII			
SeqDerMap	FrgrPrnt	SeqDerMap	FrgrPrnt	SeqDerMap	FrgrPrnt
8696	8777	4710	4671	2282	2251
6	<800	6382	6514	2067	2068
8311	8264	512	<800	5626	5794
3794	3847	449	<800	4289	4117
517	<800	1417	1461	4382	4412
5654	5717	5366	5126	4594	4580
3530	3552	590	<800	10637	10624
2635	2737	3606	3651	4577	4580
3532	3552	274	<800	246	<800
12069	11650	4570	4539	933	948
8046	8264	9049	8995	15529	15666
2974	3009	4374	4320	184	<800
3451	3552	4978	5126	4296	4283
2090	2126	11396	11164	2736	2818
1581	1500	1746	1719	3990	3991
555	<800	1599	1541	4246	4117
210	<800	8101	8204	693	<800
1074	1051	2394	2339	3597	3667
2455	2540	1944	1928	12762	12592
8953	8777	5118	5126	1327	1358
361	<800	3069	3110	3651	3667
3186	3259	2276	2339	1831	1785
19747	19152	7388	7507	404	<800
1669	1668	196	<800	3353	3348
1367	1329	312	<800	159	<800
3124	3171	2294	2339	4303	4283
24350	24894	5185	5126	483	<800
6070	6075	714	<800	4395	4412
1385	1329	1355	1358	6197	6227

FEATURES				Location/Qualifiers			
source				1..182509			
Query Match				67.7%; Score 2619.6; DB 9; Length 182509;			
Best Local Similarity				99.8%; Pred No. 0;			
Matches 2633; Conservative				0; Mismatches 4; Indels 1; Gaps 1;			
Qy	1213	AGGGAGAGTTGTGGAATATGTGGATGACCTCTTTGGAACTGGAGGAGACAGCTAGCCCA	1272				
Db	178281	AGGGAGAGTTGTGGAATATGTGGATGACCTCTTTGGAACTGGAGGAGACAGCTAGCCCA	178340				
Qy	1273	CAGCAACCAAGAGACTTCTCTTGGCGTTCAGGAACACAGATCTTGTCTTTTCCC	1332				
Db	178341	CAGCAACCAAGAGACTTCTCTTGGCGTTCAGGAACACAGATCTTGTCTTTTCCC	178400				
Qy	1333	AACGCCAGGCTGTGTATACCTCAGAGCCCTCTCTTTACTCTCCAAAGTGAAGGGAAG	1392				
Db	178401	AACGCCAGGCTGTGTATACCTCAGAGCCCTCTCTTTACTCTCCAAAGTGAAGGGAAG	178460				

Qy	1393	CCCGGCTCTCTAACTGATGTATCAGGGGTGAGCTGCGCTTTCTCTATCTTCACACCT	1452				
Db	178461	CCCGGCTCTCTAACTGATGTATCAGGGGTGAGCTGCGCTTTCTCTATCTTCACACCT	178520				
Qy	1453	GCACCTCATGTTTACACCTATCTTTCTCACCTTTTGTGAGATGGAGTCTCGCTCTCT	1512				
Db	178521	GCACCTCATGTTTACACCTATCTTTCTCACCTTTTGTGAGATGGAGTCTCGCTCTCT	178580				
Qy	1513	TGCCAGGCTGGAGTGAATGGACAGCTTCTCAGCTCAGTGCACCTCCGCTCTTGGTT	1572				
Db	178581	TGCCAGGCTGGAGTGAATGGACAGCTTCTCAGCTCAGTGCACCTCCGCTCTTGGTT	178640				
Qy	1573	CAAGCAATTCGTGTCATCAGCTCCGAGTACTTGGGATTACAGGCATGTGCACCAAG	1632				
Db	178641	CAAGCAATTCGTGTCATCAGCTCCGAGTACTTGGGATTACAGGCATGTGCACCAAG	178700				
Qy	1633	CCCGGCTAAATTTGTATTTTGTAGTAGAGAGGGGTTTTCGCACTGTGGCCAGGCTGTCT	1692				
Db	178701	CCCGGCTAAATTTGTATTTTGTAGTAGAGAGGGGTTTTCGCACTGTGGCCAGGCTGTCT	178760				
Qy	1693	CGAACTCTTTGACTTCAGATGATCCATCTGCTTGGCTCCACAGTGTGGGATTACAGG	1752				
Db	178761	CGAACTCTTTGACTTCAGATGATCCATCTGCTTGGCTCCACAGTGTGGGATTACAGG	178820				
Qy	1753	CGTAGCCACCATGCCCCGCTCTTTCTCACCTTTACACTGCTTCTTATCTCATCATC	1812				
Db	178821	CGTAGCCACCATGCCCCGCTCTTTCTCACCTTTACACTGCTTCTTATCTCATCATC	178880				
Qy	1813	TGTTTTACACCTTCATCT	1872				
Db	178881	TGTTTTACACCTTCATCT	178940				
Qy	1873	GCTGCTTCTTACCAATTTGGTTTGAAGGCGAGTCTCTCTGGCTGTGTTTTTTTTT	1932				
Db	178941	GCTGCTTCTTACCAATTTGGTTTGAAGGCGAGTCTCTCTGGCTGTGTTTTTTTTT	179000				
Qy	1933	TCCAGAAATCAGTATTATTTTAAATAAGAAACATCTCTAGAGATGATATTCT	1992				
Db	179001	TCCAGAAATCAGTATTATTTTAAATAAGAAACATCTCTAGAGATGATATTCT	179060				
Qy	1993	GAAACCTCTCTTGGCTTATTTGCTTTTCCAGATTTTAGTCTCTCTCTCTCTCTCTCT	2052				
Db	179061	GAAACCTCTCTTGGCTTATTTGCTTTTCCAGATTTTAGTCTCTCTCTCTCTCTCTCT	179120				
Qy	2053	AAAGATGGTGAAGACATAGGCTAAATTTCTCGAGCTCACAATGCTCTCTCTCTCTCT	2112				
Db	179121	AAAGATGGTGAAGACATAGGCTAAATTTCTCGAGCTCACAATGCTCTCTCTCTCTCT	179180				
Qy	2113	GACTTGTACCAATTCAGACCCACTGAAAAACAAGTTGAGTAGAGTGTAGAGTGCAG	2172				
Db	179181	GACTTGTACCAATTCAGACCCACTGAAAAACAAGTTGAGTAGAGTGTAGAGTGCAG	179240				
Qy	2173	AAATGGCTTTTGGCCCACTTTGTCATCTCCAAAAATTAACAAGTTGGCCGATGCCATTT	2232				
Db	179241	AAATGGCTTTTGGCCCACTTTGTCATCTCCAAAAATTAACAAGTTGGCCGATGCCATTT	179300				
Qy	2233	GAGGCAATCTTAGTTATAAGTCTCCGAGTTTGAAGGAGAAAGCAGAGCTGTCTA	2292				
Db	179301	GAGGCAATCTTAGTTATAAGTCTCCGAGTTTGAAGGAGAAAGCAGAGCTGTCTA	179360				
Qy	2293	GTTCATTCATTTCTTTCAGTAAATATTTATTTAGTACCTTACTGTGTGTAGGCAATTGACC	2352				
Db	179361	GTTCATTCATTTCTTTCAGTAAATATTTATTTAGTACCTTACTGTGTGTAGGCAATTGACC	179420				
Qy	2353	TGGGAATCAGATACCTTCAGAGATACAGGGAAGTTCCCTGTCTCTAGGAGCTTAC	2412				
Db	179421	TGGGAATCAGATACCTTCAGAGATACAGGGAAGTTCCCTGTCTCTAGGAGCTTAC	179480				
Qy	2413	ATTCTACGGGAGAAAGATAGCCAAATACATAGGAATAAATATATACAGGTATCATGT	2472				
Db	179481	ATTCTACGGGAGAAAGATAGCCAAATACATAGGAATAAATATATACAGGTATCATGT	179540				

```
QY 2473 AGTGATTAATGCTGTCGAGAAATAAAGCAGGGAGGAGTAAGAAATCCTGGAGATGA 2532
Db 179541 AGTGATTAATGCTGTCGAGAAATAAAGCAGGGAGGAGTAAGAAATCCTGGAGATGA 179600
QY 2533 GGCTGCAGTCTTTAAATGGGGCTCTCACTGGGAATGTGACCTTCAGCAGAGACGCTTAGGAA 2592
Db 179601 GGCTGCAGTCTTTAAATGGGGCTCTCACTGGGAATGTGACCTTCAGCAGAGACGCTTAGGAA 179660
QY 2593 GTGGATCCTGGCAAGGCATTCACGACAGAGAAACAAGATGTGCACTGCCCCAAAGTCAG 2652
Db 179661 GTGGATCCTGGCAAGGCATTCACGACAGAGAAACAAGATGTGCACTGCCCCAAAGTCAG 179720
QY 2653 AACTCTCTCTACGTGTCTAGGAAGAGCAGAGAGCAACAGCAGAGTCGTGGCAGGGCTA 2712
Db 179721 AACTCTCTCTACGTGTCTAGGAAGAGCAGAGAGCAACAGCAGAGTCGTGGCAGGGCTA 179780
QY 2713 GAATGAAAGAGAGGCGCTGGGAGGACAGATGTGTGGAGGCGCTTGGCTTCTGCTAAGT 2772
Db 179781 GAATGAAAGAGAGGCGCTGGGAGGACAGATGTGTGGAGGCGCTTGGCTTCTGCTAAGT 179840
QY 2773 GAGATGGGAACCACTGGAGGCTTGAACAGAGAGTCCTTGATTTATATTTTGA 2832
Db 179841 GAGATGGGAACCACTGGAGGCTTGAACAGAGAGTCCTTGATTTATATTTTGA 179900
QY 2833 AGGTCATCTAGCTGCAATATTTGTGAATACTTTAGTGGCAAGGGCAGAGGAGG 2892
Db 179901 AGGTCATCTAGCTGCAATATTTGTGAATACTTTAGTGGCAAGGGCAGAGGAGG 179960
QY 2893 GAAGACCTGTAGGAAGTACTGCAAGGTTCCAGGTTCCAGGCTTGGGCTGGGCCAGCAACAGC 2952
Db 179961 GAAGACCTGTAGGAAGTACTGCAAGGTTCCAGGTTCCAGGCTTGGGCTGGGCCAGCAACAGC 180020
QY 2953 AGTGTGCAATATCTAGATTTTATTTGAAAGAGCAATAGGATTTGCTGAGAGTTGAA 3012
Db 180021 AGTGTGCAATATCTAGATTTTATTTGAAAGAGCAATAGGATTTGCTGAGAGTTGAA 180080
QY 3013 TGTGAGTGTAGAGAGGAGTAAATGATGAATTAAGTATTTTGGCTGCAATAGCA 3072
Db 180081 TGTGAGTGTAGAGAGGAGTAAATGATGAATTAAGTATTTTGGCTGCAATAGCA 180140
QY 3073 GGAAGATGGATTACCTACTGTAATAGGAGGAGTGGCTGGGTAACTAGGATTT 3132
Db 180141 GGAAGATGGATTACCTACTGTAATAGGAGGAGTGGCTGGGTAACTAGGATTT 180200
QY 3133 TGTGCAAGAGCGGTCTGTGTTGGTAATGGAGGTTCTGGCTGCAAAATCAAGTGG 3192
Db 180201 TGTGCAAGAGCGGTCTGTGTTGGTAATGGAGGTTCTGGCTGCAAAATCAAGTGG 180260
QY 3193 GA-TTCTCTCAGTCAGTCTGACAGAGCTCGAGACAGGATCTGAATGCACTTGGTT 3251
Db 180261 GAGTCTCTCAGTCAGTCTGACAGAGCTCGAGACAGGATCTGAATGCACTTGGTT 180320
QY 3252 TATTTGTTGGGGTGTCTCTCAGAGGAACTGTGAAAGGCTTTATCAGTCATTTATTGGCT 3311
Db 180321 TATTTGTTGGGGTGTCTCTCAGAGGAACTGTGAAAGGCTTTATCAGTCATTTATTGGCT 180380
QY 3312 GTGAGAGTCTCTGGAGTGTGGTCAATTTGAAGGCAAGTGAATTCAGTTGAGGCA 3371
Db 180381 GTGAGAGTCTCTGGAGTGTGGTCAATTTGAAGGCAAGTGAATTCAGTTGAGGCA 180440
QY 3372 GTCTCTGAAAAAGAGCTGTAGGCATCTGGCAGCTACCATGCTGGTGTAGTGTGGGG 3431
Db 180441 GTCTCTGAAAAAGAGCTGTAGGCATCTGGCAGCTACCATGCTGGTGTAGTGTGGGG 180500
QY 3432 TGGGGTCTCTGGGCACTGCTGTGTGAAGGATCTGGCAGGCAACACAGCGCCCCCTAC 3491
Db 180501 TGGGGTCTCTGGGCACTGCTGTGTGAAGGATCTGGCAGGCAACACAGCGCCCCCTAC 180560
QY 3492 TGAAACCATCAGCATGTCAGTGCATTTAAAGCCATGCACTGGAGGGCCACTGATTTG 3551
Db 180561 TGAAACCATCAGCATGTCAGTGCATTTAAAGCCATGCACTGGAGGGCCACTGATTTG 180620
QY 3552 TCTCTGAGTATTACTGAGAAGCAACAGAAAAAGAGCCATGGATGGAGGCCCTTGGGCTCTCT 3611
```

```
Db 180621 TCTCTGAGTATTACTGAGAAGCAACAGAAAAAGAGCCATGGATGGAGGCCCTTGGGCTCTCT 180680
QY 3612 GGGAAATGGGAATCAGCCAAAGGACTGGAAGAGGTTACCTTAAGGTGAGAAAAACCA 3671
Db 180681 GGGAAATGGGAATCAGCCAAAGGACTGGAAGAGGTTACCTTAAGGTGAGAAAAACCA 180740
QY 3672 AGAGAGTGTGGTGTCTTGGAGAGCTGAGCTTTCTTTTATTTCAACCTCATTTCCCTTCTCCAA 3731
Db 180741 AGAGAGTGTGGTGTCTTGGAGAGCTGAGCTTTCTTTTATTTCAACCTCATTTCCCTTCTCCAA 180800
QY 3732 TAAGCCACTTGTAGTGTGGGCCCCCTCCAGGTTTGAAGCAAGAGGAGAAAGCAGCG 3791
Db 180801 TAAGCCACTTGTAGTGTGGGCCCCCTCCAGGTTTGAAGCAAGAGGAGAAAGCAGCG 180860
QY 3792 TTTGGGAAACAAGACACTTTTCTGCAATAGCTCGGAGGAATAAAAAAGGATAGAGTGT 3849
Db 180861 TTTGGGAAACAAGACACTTTTCTGCAATAGCTCGGAGGAATAAAAAAGGATAGAGTGT 180918
```

RESULT 3
AC123900
LOCUS
DEFINITION Homo sapiens chromosome 3 clone RP11-640D6, linear PRI 29-JUN-2002
AC123900
ACCESSION
VERSION
AC123900.2 GI:21629387
KEYWORDS
HIG.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Chordata; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 166485)
Kaul R.K., Olson M.V., Zhou Y., James R.A., Rouse G., Wu Z.,
Saenphimmachak C., Phelps K.A., Buckley D., Kibukawa M., Raymond C.,
and Haugen B.D.
Direct Submission
Unpublished
2 (bases 1 to 166485)
Kaul R.K., Olson M.V., Raymond C. and Haugen B.D.
Direct Submission
Submitted (04-JUN-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
3 (bases 1 to 166485)
Kaul R.K., Olson M.V., Zhou Y., James R.A., Rouse G., Wu Z.,
Saenphimmachak C., Phelps K.A., Buckley D., Kibukawa M., Raymond C.,
and Haugen B.D.
Direct Submission
Submitted (29-JUN-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
On Jun 29, 2002 this sequence version replaced gi:21321810.
----- Genome Center
Center: University of Washington Genome Center
Center Code: UWGC
Web site: <http://www.genome.washington.edu>
Contact: uwgctgs@u.washington.edu
----- Project Information
Center project name: chr-3
Center clone name: RP11-640D6 (bc0506)
----- Summary Statistics
Sequencing vector: plasmid; 100% of reads
Chemistry: Dye-terminator ET; 19% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 166343 bases at least Q40
Consensus quality: 166451 bases at least Q30
Consensus quality: 166480 bases at least Q20
Insert size: 166485; sum-of-contigs
Quality coverage: 7.5x in Q20 bases; sum-of-contigs

Overlapping Sequences:
5': RP11-627J17 (UWGC:bc0502) AC112211 21818-bp overlap
3': RP11-294018 (UWGC:bc0346) AC122176

Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program.

All manually edited bases have been reduced to quality zero.

Quality levels above 40 are expected to have less than 1 error in 10,000 bp.

Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:

This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

EcoRI

BglII

HindIII

SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt
8696	8737	2811	2816	1461	1410
6	<800	2067	2044	6382	6450
3530	3515	11749	11492	512	<800
5654	5722	4594	4513	449	<800
517	<800	4382	4513	8031	8087
3794	3785	4295	4513	4570	4537
8317	8204	2621	2816	274	<800
1659	1620	1611	1564	3606	3653
8141	8204	1043	1015	590	<800
646	<800	3299	3403	5372	5286
2223	2226	885	861	3079	3054
3811	3785	9258	9095	775	<800
3020	3008	282	<800	659	<800
2953	3008	2006	2044	1190	1170
3240	3226	11003	11492	1367	1328
3763	3785	8	<800	4836	4684
13647	13653	4619	4513	408	<800
3457	3515	449	<800	21	<800
379	<800	1955	2044	258	<800

FEATURES

Location/Qualifiers

10453	10224	2840	2816	4524	4537
5001	5031	1561	1564	3411	3436
388	<800	6877	6790	948	911
832	842	16065	15540	272	<800
165	<800	4076	3837	4359	4227
116	<800	15546	15540	210	<800
4322	4241	1520	1564	4666	4684
6490	6601	141	<800	854	911
3319	3327	83	<800	10084	10002
1081	1076	7283	7183	1420	1410
7379	7393	844	861	2040	2051
907	918	3456	3403	2972	3054
5434	5406	412	<800	4263	4227
14267	14476	4550	4513	4172	4227
7852	7869	439	<800	386	<800
2483	2526	12228	11492	3588	3653
1573	1533	525	<800	267	<800
2706	2706	2095	2044	10609	10500
5788	5722	533	<800	11893	11650
1413	1381	386	<800	645	<800
12127	11968	308	<800	1350	1328
901	918	2561	2598	1765	1742
2140	2117	3867	3837	1002	1007
430	<800	11514	11492	1619	1590
254	<800	743	731	4250	4227
		1621	1564	15992	16197
		720	731	5279	5286
		2767	2816	3965	4227
		144	<800	4595	4684
		542	<800	5928	5910
				900	911
				6373	6450
				1882	1867
				539	<800
				559	<800
				3763	3959

source	1. 166495 /organism="Homo sapiens" /mol_type="genomic DNA"	
Query Match	67.3%; Score 2605.2; DB 9; Length 166485;	
Best Local Similarity	99.6%; Pred. No. 0;	
Matches 2634; Conservative	0; Mismatches 3; Indels 7; Gaps 2;	
QY	1213 AGGAGAAAGTTGGGAATATGTGGATGACCTCTTGGAACTGGAGGAGACAGCTAGCCCA	1272
DB	17584 AGGAGAAAGTTGGGAATATGTGGATGACCTCTTGGAACTGGAGGAGACAGCTAGCCCA	17643
QY	1273 CAGCAACCAAGAGACTTCCTCTTGGGCTTCAGGAACACAGATCTCTTGTCTCTTCC	1332
DB	17644 CAGCAACCAAGAGACTTCCTCTTGGGCTTCAGGAACACAGATCTCTTGTCTCTTCC	17703
QY	1333 AACAGCCAGCTGTGTGATACCTCAGAGCTTCTCTTTACTCTCCAAAGTGAAGGAAG	1392
DB	17704 AACAGCCAGCTGTGTGATACCTCAGAGCTTCTCTTTACTCTCCAAAGTGAAGGAAG	17763
QY	1393 CCCCCGTCTCTAACTGCATGTCATCAGGGGTGAGCTGCTTCTCTATCTTCACACT	1452
DB	17764 CCCCCGTCTCTAACTGCATGTCATCAGGGGTGAGCTGCTTCTCTATCTTCACACT	17823
QY	1453 GCACCTCATGTTTCACACCTATCTTCTCACCTTTTTTTTGTGATGGAGTCTCGCTCT	1512
DB	17824 GCACCTCATGTTTCACACCTATCTTCTCACCTTTTTTTTGTGATGGAGTCTCGCTCT	17883
QY	1513 TGCCAGGCTGGAGTGCATGSCAGCTTCTCAGCTCAGTGCAGAACTCCGCTCTTGGGT	1572
DB	17884 TGCCAGGCTGGAGTGCATGSCAGCTTCTCAGCTCAGTGCAGAACTCCGCTCTTGGGT	17943
QY	1573 CAAGCAATCTGCTGATCAGCTCCCGAGTACCTGGGATACAGGCATGTGCCACACAG	1632
DB	17944 CAAGCAATCTGCTGATCAGCTCCCGAGTACCTGGGATACAGGCATGTGCCACACAG	18003
QY	1633 CCGGCTAAATTTGTATTTTAGTAGAGACGGGTTTGGCATGTGGCAGGCTGGTCT	1692
DB	18004 CCGGCTAAATTTGTATTTTAGTAGAGACGGGTTTGGCATGTGGCAGGCTGGTCT	18063
QY	1693 CGAATCTTGTGACTTCAGATGATCCATCTGCTTGGCTCCACAGTGTGGGATACAGG	1752
DB	18064 CGAATCTTGTGACTTCAGATGATCCATCTGCTTGGCTCCACAGTGTGGGATACAGG	18123
QY	1753 CGTAGCCACCATGCGCGCTCTTCTCACCTTTACACCTGTCTTTATCTCCACATC	1812
DB	18124 CGTAGCCACCATGCGCGCTCTTCTCACCTTTACACCTGTCTTTATCTCCACATC	18183
QY	1813 TGTFTTTCACACCTTCATCCCTCTCTCTCATGTTTCAGCTGTCTTCCCATGTTTATA	1872
DB	18184 TGTFTTTCACACCTTCATCCCTCTCTCTCATGTTTCAGCTGTCTTCCCATGTTTATA	18243
QY	1873 GCTGCTTTCTTACCATTTTGGTTTGAAGGCGAGTCTTCTCTGGCTTGTTTTTTGT	1932
DB	18244 GCTGCTTTCTTACCATTTTGGTTTGAAGGCGAGTCTTCTCTGGCTTGTTTTTTGT	18303
QY	1933 TCCCGAAATCAGTATATTTTTTAAATAAGAAACATCTCTAGAGATGATAATGT	1992
DB	18304 TCCCGAAATCAGTATATTTTTTAAATAAGAAACATCTCTAGAGATGATAATGT	18363
QY	1993 GAAACCTCTCTTGGCTTATTTGGCTTTCCAGATTTTAGTCTCTCTCCCATCCGGG	2052
DB	18364 GAAACCTCTCTTGGCTTATTTGGCTTTCCAGATTTTAGTCTCTCTCCCATCCGGG	18423
QY	2053 AAAGATGTTGGAAGACATAGGCTAAATTTCTCCAGCTTCACAAATGGTCTTACATGGT	2112
DB	18424 AAAGATGTTGGAAGACATAGGCTAAATTTCTCCAGCTTCACAAATGGTCTTACATGGT	18483
QY	2113 GACTTGTACCAATCTCAGACCCACTGAAACAAAGTTGAGTAGAGATGAGAGTGCAG	2172
DB	18484 GACTTGTACCAATCTCAGACCCACTGAAACAAAGTTGAGTAGAGATGAGAGTGCAG	18543
QY	2173 AAATGTGGCTTTTGGCCCACTTTGCATCTCCAAATTAACAGGTTGGCCGATCCCATTT	2232
DB	18544 AAATGTGGCTTTTGGCCCACTTTGCATCTCCAAATTAACAGGTTGGCCGATCCCATTT	18603
QY	2233 GAGGACAAATGCTTAGTTTATAAGTCTCCGAGTTCGAAAGGAAGAACCCAGAGCTGTCTA	2292
DB	18604 GAGGACAAATGCTTAGTTTATAAGTCTCCGAGTTCGAAAGGAAGAACCCAGAGCTGTCTA	18663
QY	2293 GTTTCATTCATCTTCTCAGTAAATATTTTATGTAGTACCTACTGTGTGCTAGGATTGACC	2352
DB	18664 GTTTCATTCATCTTCTCAGTAAATATTTTATGTAGTACCTACTGTGTGCTAGGATTGACC	18723
QY	2353 TGG-----GAATAGAGATACCTCAGAAATACAGGGAAGTTCCTCTGTCTCATGGA	2406
DB	18724 TGGGAATCAGAACTAGAGATACCTCAGAAATACAGGGAAGTTCCTCTGTCTCATGGA	18783
QY	2407 GTTTACATTTCTACAGGGAAGAGATAGCAATATACATAGGAATAATATATCAAGGTA	2466
DB	18784 GTTTACATTTCTACAGGGAAGAGATAGCAATATACATAGGAATAATATATCAAGGTA	18843
QY	2467 TCATGTAGTATTAATTTGCTGTGGAGAAATTAAGCAGGGGGAGTAAAGTAATCTCTGG	2526
DB	18844 TCATGTAGTATTAATTTGCTGTGGAGAAATTAAGCAGGGGGAGTAAAGTAATCTCTGG	18903
QY	2527 AGATGAGGCTGCAGTTTAAATGGGCTCCTCCTGGGAATGTGACGTTGAGCAGAGAGCTT	2586
DB	18904 AGATGAGGCTGCAGTTTAAATGGGCTCCTCCTGGGAATGTGACGTTGAGCAGAGAGCTT	18963
QY	2587 AGGGAAGTGGATCTCGACAGAGCATCTCAGGCGAGGAAACAGATGTGCACCTGCCCAA	2646
DB	18964 AGGGAAGTGGATCTCGACAGAGCATCTCAGGCGAGGAAACAGATGTGCACCTGCCCAA	19023
QY	2647 AGTGAGAACTTGTCTACGTGGTTCAGGAAGAGCAGGAGACCAAGCAGAGTCTGGGCA	2706
DB	19024 AGTGAGAACTTGTCTACGTGGTTCAGGAAGAGCAGGAGACCAAGCAGAGTCTGGGCA	19083
QY	2707 GGGTAGAATGGAAGAGAGGCGCTGGGGAGGACAGGTGTGGAGGGCTTGGCTTCTG	2766
DB	19084 GGGTAGAATGGAAGAGAGGCGCTGGGGAGGACAGGTGTGGAGGGCTTGGCTTCTG	19143
QY	2767 CTAACTGAGATGGAACTTCTGAGGAGTTCGAGAGGAGTTCGCTTGTATTTATAT	2826
DB	19144 CTAACTGAGATGGAACTTCTGAGGAGTTCGAGAGGAGTTCGCTTGTATTTATAT	19203
QY	2827 TTTGCAAGGCTCATTTAGCTGCAATATTTGTGAAAACTTTTGTGACAAAGGCGAGAAG	2886
DB	19204 TTTGCAAGGCTCATTTAGCTGCAATATTTGTGAAAACTTTTGTGACAAAGGCGAGAAG	19263
QY	2887 AAGGGAAGACCTGTTAGGAGCTACTGCAAGGTTCCAGGTTTGGGCTGGCCACAGC	2946
DB	19264 AAGGGAAGACCTGTTAGGAGCTACTGCAAGGTTTGGGCTGGCCACAGC	19323
QY	2947 AACAGCAGTGGTCAAAATATCTAGATTTATTTTGAAGAGCAATAGGATTTGCTGAGAG	3006
DB	19324 AACAGCAGTGGTCAAAATATCTAGATTTATTTTGAAGAGCAATAGGATTTGCTGAGAG	19383
QY	3007 TTTGAATGTGGAGTGTAAAGAAAGAGAGTTAATGATGACATTAAAGTTTTGGCTGA	3066
DB	19384 TTTGAATGTGGAGTGTAAAGAAAGAGAGTTAATGATGACATTAAAGTTTTGGCTGA	19443
QY	3067 ATAGCAGGAAAGATGGAGTTTACAGTTACTGAAATAGGGAAGGATGGGCTGGGTAAATA	3126
DB	19444 ATAGCAGGAAAGATGGAGTTTACAGTTACTGAAATAGGGAAGGATGGGCTGGGTAAATA	19503
QY	3127 GGAATTTGGTGAAGACAGGCTGTCTGTGGTGGAAATGGAGGTTCTGGCTGCAATCAA	3186
DB	19504 GGAATTTGGTGAAGACAGGCTGTCTGTGGTGGAAATGGAGGTTCTGGCTGCAATCAA	19563
QY	3187 AGTGGAGA-TTCTCTCAGGCTGAGGCTGAGCAGAGCTCGACACAGGATCTGAATGCAC	3245
DB	19564 AGTGGAGA-TTCTCTCAGGCTGAGGCTGAGCAGAGCTCGACACAGGATCTGAATGCAC	19623
QY	3246 TTGGTTTATTTGGGGGTGCTCTCAGAGGAACCTGTGAAAGCCCTTATCAGTCAATTTA	3305

D _b	602	CATTAAAGACTTGGAAACCAAGTCATATGAAAGCGCTGTTCATCGAGCAGCTGCCGGCATA	661
Q _y	713	CAAACGGTGAGAACCTGGGAACAATCCATCACAGACATGGAGCTGGCCCTTCCCGACTTCCT	772
D _b	662	CAAACGGTGAGAACCTGGGAACAATCCATCACAGACATGGAGCTGGCCCTTCCCGACTTCCT	721
Q _y	773	CAAAGCCCTTTTACGAGTGTCTCGAGCCTCGAGGGTTCCAGGAGATCAAGACCTTCAA	832
D _b	722	CAAAGCCCTTTTACGAGTGTCTCGAGCCTCGAGGGTTCCAGGAGATCAAGACCTTCAA	781
Q _y	833	GGATTTCTACCTTTCCATAGCAGATCATATGTAGAAGTTCTGGAATGC AAAATACAGTG	892
D _b	782	GGATTTCTACCTTTCCATAGCAGATCATATGTAGAAGTTCTGGAATGC AAAATACAGTG	841
Q _y	893	TGAAGAGAACCTCAACCCAGATTATAGAGGCTATCCCGTTGAGAAAATTTGFGGTACCAAT	952
D _b	842	TGAAGAGAACCTCAACCCAGATTATAGAGGCTATCCCGTTGAGAAAATTTGFGGTACCAAT	901
Q _y	953	GTAATCATTTACTTGCAGTTTGCTATATAAGTTCGAACGCTGAAGAAATGCAGCCCCCTG	1012
D _b	902	GTAATCATTTACTTGCAGTTTGCTATATAAGTTCGAACGCTGAAGAAATGCAGCCCCCTG	961
Q _y	1013	TGCACTCAGCTATCTGCTCTTTGATCAGAAATGCAAGGTCATGCAGCAGAACCTGTGTATA	1072
D _b	962	TGCACTCAGCTATCTGCTCTTTGATCAGAAATGCAAGGTCATGCAGCAGAACCTGTGTATA	1021
Q _y	1073	TTACCAAGTACCACAGGAGACATTTGGGCCCTCTCGATGAGACATTCAGGCCAGACCTGA	1132
D _b	1022	TTACCAAGTACCACAGGAGACATTTGGGCCCTCTCGATGAGACATTCAGGCCAGACCTGA	1081
Q _y	1133	AGCAGTTTCAGTCTTTAATGTGACCAACACTCCAGAGGAGCTGTATGACTTTGCTAAGGA	1192
D _b	1082	AGCAGTTTCAGTCTTTAATGTGACCAACACTCCAGAGGAGCTGTATGACTTTGCTAAGGA	1141
Q _y	1193	AAATATAATGAGATGATGATAGGAGGAAGTTGTGGAATATGTGATGACCTCTTGGAACCT	1252
D _b	1142	AAATATAATGAGATGATGATAGGAGGAAGTTGTGGAATATGTGATGACCTCTTGGAACCT	1201
Q _y	1253	GGAGGAGACAGCTTAGCCACAGCAACCAAGAGACATCTCTTGGCGGTTTCAGGAACAC	1312
D _b	1202	GGAGGAGACAGCTTAGCCACAGCAACCAAGAGACATCTCTTGGCGGTTTCAGGAACAC	1261
Q _y	1313	AGATTCCTTTGFCCTTTTCCC AACAGCCAGGCTCTTGATACCTCAGAGCCTTCTCTTTTAC	1372
D _b	1262	AGATTCCTTTGFCCTTTTCCC AACAGCCAGGCTCTTGATACCTCAGAGCCTTCTCTTTTAC	1321
Q _y	1373	TCTCCAAGTGAAAGGAGAACCCCGCTCTCTTAACCTGATGATCAGGGGTGAGCGTG	1432
D _b	1322	TCTCCAAGTGAAAGGAGAACCCCGCTCTCTTAACCTGATGATCAGGGGTGAGCGTG	1381
Q _y	1433	CCTTTCCTATCTTCACACCTGCCACCTCATGTTCCACACCTATCTTCTCACCTTTTTTTTT	1492
D _b	1382	CCTTTCCTATCTTCACACCTGCCACCTCATGTTCCACACCTATCTTCTCACCTTTTTTTTT	1441
Q _y	1493	GAGATGAGTCTCGCTCTCTTGCCAGGCTGGAGTGAATGGACGTTCTCAGCTCACTG	1552
D _b	1442	GAGATGAGTCTCGCTCTCTTGCCAGGCTGGAGTGAATGGACGTTCTCAGCTCACTG	1501
Q _y	1553	CAACCTCCGCTCTTGGGTTTCAAGCAATTTCTGTGCATCAGCCTCCCGAGTACTCGGAT	1612
D _b	1502	CAACCTCCGCTCTTGGGTTTCAAGCAATTTCTGTGCATCAGCCTCCCGAGTACTCGGAT	1561
Q _y	1613	TACAGGCAATGTCACCAACGCCCCGGCTAAATTTGTAATTTTAA -- GTAGACACGGGTTTT	1670
D _b	1562	TACAGGCAATGTCACCAACGCCCCGGCTAAATTTGTAATTTTAA -- GTAGACACGGGTTTT	1621
Q _y	1671	GCCATGTTGGCAGGCTGTTCTGAACTCTTGACTTCAGATGATCCATCTGCGCTT -- GGCC	1729
D _b	1622	GCCATGTTGGCAGGCTGTTCTGAACTCTTGACTTCAGATGATCCATCTGCGCTT -- GGCC	1681
Q _y	1730	TCCACAGTCTGGGATTTACAGGCGTGAGCCACCATGCCCCGGCTCTTTCTCACCTTTTAC	1789

Db	1682	TCCACAGTGTGGGATTACAGCGGTGAGCCACCATCGCCGGCTCTTTCTCACCTTTAC	1741
Qy	1790	ACCTGTCTTCTTATCCTCACAATCTGTTTTTCCACACCTTTCATCCTGTCTCTCTCATGTGTA	1849
Db	1742	ACCTGTCTTCTTATCCTCACAATCTGTTTTTCCACACCTTTCATCCTGTCTCTCTCATGTGTA	1801
Qy	1850	CATTGTCTTCCCAATGTTTCATAGTGCCTTTCTTACCAATTTTGTTGGTTTGAAGGCGAGTCT	1909
Db	1802	CATTGTCTTCCCAATGTTTCATAGTGCCTTTCTTACCAATTTTGTTGGTTTGAAGGCGAGTCT	1861
Qy	1910	TCTCTGGCTTGTTTTTTGTTTTTTCCAGAAAAATCAGTATTATTTTTTAAATTAAGAAAAA	1969
Db	1862	TCTCTGGCTTGTTTTTTGTTTTTTCCAGAAAAATCAGTATTATTTTTTAAATTAAGAAAAA	1921
Qy	1970	CATTCTAGAAAGATGATAATGTGAAACACTCTTTGGCTTATTGCTTTTCAGATTTT	2029
Db	1922	CATTCTCTAGAAAGATGATAATGTGAAACACTCTTTGGCTTATTGCTTTTCAGATTTT	1981
Qy	2030	AGTCTCCTTTCTCCCAATCCGGAACAGTGGTGGAGACATAGCTTAAATTTCTCCAGCC	2089
Db	1982	AGTCTCCTTTCTCCCAATCCGGAACAGTGGTGGAGACATAGCTTAAATTTCTCCAGCC	2041
Qy	2090	TCACAAATGCTTCTCACTTGGTCTGACTTGTACCAATTTAGCACCCACTGAAAAACAAGT	2149
Db	2042	TCACAAATGCTTCTCACTTGGTCTGACTTGTACCAATTTAGCACCCACTGAAAAACAAGT	2101
Qy	2150	TGAGTAGAGAGTGTAGAGTCAGAAATGTGGCTTTTGGCCCACTTTGCATCTCCAAAAT	2209
Db	2102	TGAGTAGAGAGTGTAGAGTCAGAAATGTGGCTTTTGGCCCACTTTGCATCTCCAAAAT	2161
Qy	2210	ACAACGGTTGGCCGATCCCACTTTGAGGACAAAGCTTTAGTTATAAGTCTCCGAGTTGGAAA	2269
Db	2162	ACAACGGTTGGCCGATCCCACTTTGAGGACAAAGCTTTAGTTATAAGTCTCCGAGTTGGAAA	2221
Qy	2270	AGGAAGAAAGCCAGAGCTGTCTAGTTTCATTTCAATCTTTTCAGTAATAATTTATGTAGTA	2328
Db	2222	AGGAAGAAAGCCAGAGCTGTCTAGTTTCATTTCAATCTTTTCAGTAATAATTTATGTAGTA	2280
RESULT 5			
AK075401		2010 bp	mrna linear PRI 03-SEP-2002
LOCUS	AK075401	2010 bp	mrna linear PRI 03-SEP-2002
DEFINITION	Homo sapiens cDNA PSEC0091 fis, clone PLACE1001683, highly similar to Homo sapiens mRNA for cartilage-associated protein.		
ACCESSION	AK075401		
VERSION	AK075401.1 GI:22761465		
KEYWORDS	oligo capping; fis (full insert sequence).		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Oka, T., Nishikawa, T., Suzuki, Y., Kawai-Hio, Y., Hayashi, K., Ishii, S., Saito, K., Yamamoto, J., Wakamatsu, A., Nagai, T., Nakamura, Y., Sato, H., Negahari, K., Sugano, S. and Isogai, T.		
TITLE	HRI human cDNA sequencing project		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 2010)		
AUTHORS	Isogai, T. and Yamamoto, J.		
TITLE	Direct Submission		
JOURNAL	Submitted (20-WAR-2002) Takao Isogai, Helix Research Institute, Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)		
COMMENT	HRI human cDNA sequencing project; cDNA 5' - & 3'-end one pass sequencing, clone selection and full insert sequencing; Helix Research Institute (supported by Japan Key Technology Center etc.); cDNA library construction; Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center.		
FEATURES	Location/Qualifiers		
source	1. .2010		
	/organism="Homo sapiens"		
	/mol_type="mrna"		
	/db_xref="taxon:9606"		

[illegible]

QY	204	TACCGG	CA	CGGCTG	CGAC	CAAGTA	CA	CGCG	CGAG	CAC	TGCG	CGCG	CGAG	AGCT	TGCG	CTAC	GTG	263
DB	181	TACCGG	CA	CGGCTG	CGAC	CAAGTA	CA	CGCG	CGAG	CAC	TGCG	CGCG	CGAG	AGCT	TGCG	CTAC	GTG	229
QY	264	GAGATC	AGCCTT	CGGCTG	CGAC	CA	CGGCTT	CGCG	CGAC	AGCG	AGGCTT	CTT	CGCAC	CGCAAC	323			
DB	230	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	261			
QY	324	TGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	383
DB	262	TGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	321
QY	384	CTCT	TTGG	GGGCGC	TGCT	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	443
DB	322	CTCT	TTGG	GGGCGC	TGCT	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	381
QY	444	GCTT	TC	CGCC	AGTCC	CGAC	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	503
DB	382	GCTT	TC	CGCC	AGTCC	CGAC	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	441
QY	504	TACA	AGTTC	CGTAGT	TGCTT	CTT	CAAG	CGCA	TAAT	TAAT	CTCC	CCCA	AGG	CCAT	CGCG	CT	CGCG	563
DB	442	TACA	AGTTC	CGTAGT	TGCTT	CTT	CAAG	CGCA	TAAT	TAAT	CTCC	CCCA	AGG	CCAT	CGCG	CT	CGCG	501
QY	564	GCTC	CAC	ACCTT	TCT	TACT	GAAG	CACT	CTGAT	GAC	GAAAT	TGAT	GAA	GAGAA	CAT	GGCATAT	623	
DB	502	GCTC	CAC	ACCTT	TCT	TACT	GAAG	CACT	CTGAT	GAC	GAAAT	TGAT	GAA	GAGAA	CAT	GGCATAT	561	
QY	624	TATA	AGAG	CC	TGCT	GTG	CCG	AGG	ACTAC	ATT	TAAG	AC	CTG	GAA	ACC	AGT	CATATGAA	683
DB	562	TATA	AGAG	CC	TGCT	GTG	CCG	AGG	ACTAC	ATT	TAAG	AC	CTG	GAA	ACC	AGT	CATATGAA	621
QY	684	AGC	CTGT	T	CAT	CCG	AG	CAGT	CGCG	GCAT	CA	CA	CGGT	GAGAA	CT	CGG	CAAT	743
DB	622	AGC	CTGT	T	CAT	CCG	AG	CAGT	CGCG	GCAT	CA	CA	CGGT	GAGAA	CT	CGG	CAAT	681
QY	744	GACA	TG	AG	CTG	GGC	CTT	CCCG	ACTT	CTT	CAAG	AC	CTTT	AC	AGT	GT	CTCG	803
DB	682	GACA	TG	AG	CTG	GGC	CTT	CCCG	ACTT	CTT	CAAG	AC	CTTT	AC	AGT	GT	CTCG	741
QY	804	GAG	GGT	TC	CAG	GAG	AT	CA	AGG	AT	CTT	CA	AGG	AT	CTT	CA	AGG	863
DB	742	GAG	GGT	TC	CAG	GAG	AT	CA	AGG	AT	CTT	CA	AGG	AT	CTT	CA	AGG	801
QY	864	GT	AGA	GT	CT	TG	GAAT	CG	AAAT	AC	AGT	GT	GA	GAG	AA	CT	CA	923
DB	802	GT	AGA	GT	CT	TG	GAAT	CG	AAAT	AC	AGT	GT	GA	GAG	AA	CT	CA	861
QY	924	TAT	CCG	GT	TG	AAAT	TT	TG	GGCT	AC	CA	TG	TAT	CA	T	CT	TG	983
DB	862	TAT	CCG	GT	TG	AAAT	TT	TG	GGCT	AC	CA	TG	TAT	CA	T	CT	TG	921
QY	984	TT	GA	AC	CA	CT	TG	AA	AGAT	TG	CAG	CC	CT	TG	CAGT	CA	CT	1043
DB	922	TT	GA	AC	CA	CT	TG	AA	AGAT	TG	CAG	CC	CT	TG	CAGT	CA	CT	981
QY	1044	GACA	AG	GT	CAT	G	AG	CA	GAA	CC	TG	GT	TTA	CC	AGT	AC	CA	1103
DB	982	GACA	AG	GT	CAT	G	AG	CA	GAA	CC	TG	GT	TTA	CC	AGT	AC	CA	1041
QY	1104	TC	GA	T	GAG	CA	CTT	CC	AG	CC	AG	CTT	CA	AGT	TT			

QY	1284	GAGACTTCCTCTTGGCGGTTGAGGAAACACAGATTCTTTTGTCTCTTTTCCCAACAGCCGAGG	1343
Db	1222	GAGACTTCCTCTTGGCGGTTGAGGAAACACAGATTCTTTGTCTCTTTTCCCAACAGCCGAGG	1281
QY	1344	CTGTTGATACCTCAGAGCCTTCTCTTTTACTCTCCAAAGTGAAGGGGAAGCCCGCTCTCT	1403
Db	1282	CTGTTGATACCTCAGAGCCTTCTCTTTTACTCTCCAAAGTGAAGGGGAAGCCCGCTCTCT	1341
QY	1404	CTAATCGATGTCATCAGGGGTGAGCGCTGCTTTTCCTATCTTCCACCTGCCACCTCATG	1463
Db	1342	CTAATCGATGTCATCAGGGGTGAGCGCTGCTTTTCCTATCTTCCACCTGCCACCTCATG	1401
QY	1464	TTCCACCTATCTTTCTCACCTTTTTTTTCAGATGGAGTCTCGCTCTCTTTGCCAGGCGTG	1523
Db	1402	TTCCACCTATCTTTCTCACCTTTTTTTTCAGATGGAGTCTCGCTCTCTTTGCCAGGCGTG	1461
QY	1524	GAGTCAATGGCAGCTTCTCAGCTCACTCAACCTCGCGCTCTTGGGTTCGAACAAATTTCT	1583
Db	1462	GAGTCAATGGCAGCTTCTCAGCTCACTCAACCTCGCGCTCTTGGGTTCGAACAAATTTCT	1521
QY	1584	GCTGCATCAGCCTCCCGAGTACCTGGGATTACAGGCATGTGCCACACGCCCGGCTAATT	1643
Db	1522	GCTGCATCAGCCTCCCGAGTACCTGGGATTACAGGCATGTGCCACACGCCCGGCTAATT	1581
QY	1644	TTGTATTTTGTAGAGACGGGGTTTTTGGCCATGTTGGCCACGGCTGGTCTCGAACTCTTGA	1703
Db	1582	TTGTATTTTGTAGAGACGGGGTTTTTGGCCATGTTGGCCACGGCTGGTCTCGAACTCTTGA	1641
QY	1704	CTTCAGATGATCCATCTGCGCTTGGCTCCACAGTGTCTGGGATTACAGGCGTGAGCCACC	1763
Db	1642	CTTCAGATGATCCATCTGCGCTTGGCTCCACAGTGTCTGGGATTACAGGCGTGAGCCACC	1701
QY	1764	ATGCCCGGCGCTCTTCTCACCTTTTACACCTGTCTCTTATCTCTCACATCTGTGTTTTCACAC	1823
Db	1702	ATGCCCGGCGCTCTTCTCACCTTTTACACCTGTCTCTTATCTCTCACATCTGTGTTTTCACAC	1761
QY	1824	CTTCATCCCTGTCTCTCTCATGTTTCACATTTGCTTCTCCCATGTTTCATAGCTGCTTTTCT	1883
Db	1762	CTTCATCCCTGTCTCTCTCATGTTTCACATTTGCTTCTCCCATGTTTCATAGCTGCTTTTCT	1821
QY	1884	TACCATTTTGGTTGAGGGCAGTCTCTCTGCTGCTGTTTTTGTTTTTTCCAGAGAAAT	1943
Db	1822	TACCATTTTGGTTGAGGGCAGTCTCTCTGCTGCTGTTTTTGTTTTTTCCAGAGAAAT	1881
QY	1944	CAGTATTATTTTTTAAATAAGAAAAACATTCTCTAGAGATG	1984
Db	1882	CAGTATTATTTTTTAAATAAGAAAAACATTCTCTAGAGATG	1922
RESULT 9			
LOCUS	AK092990	1903 bp	linear
DEFINITION	Homo sapiens cDNA FLJ35671 fis, clone SPLEN2018180, highly similar to Homo sapiens mRNA for cartilage-associated protein.		
ACCESSION	AK092990		
VERSION	AK092990.1	GI:21751708	
KEYWORDS	oligo capping; fis (full insert sequence).		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Tashiro,H., Yamazaki,M., Watanabe,K., Kunagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Ota,T., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagaetsuna,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A., Oshina,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Nasuno,Y., Nagai,K. and Isogai,T.		
TITLE	NEDO human cDNA sequencing project		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1903)		


```

1689 CTTTACACCTGCTCTTCTTATCCTCACAATGCTTTTTCACAGCTTCATCCCTGCTCTCTCA 1748
1844 TGTTACACTTGTCTTCCCATGTTCTAGTGTGCTTCTTACCATTTTGTGAGGG 1903
1749 TGTTACACTTGTCTTCCCATGTTCTAGTGTGCTTCTTACCATTTTGTGAGGG 1808
1804 CAGTCTTCTGCTGCTGTTTTTTTGTGTTTTTCCAGAAAATCAGTATTATTTTAAATAA 1963
1809 CAGTCTTCTGCTGCTGTTTTTTTGTGTTTTTCCAGAAAATCAGTATTATTTTAAATAA 1868
1964 GAAAACATTTCTAGAGATGATTAATGTGAAA 1997
1869 GAAAACATTTCTAGAGATGAAAAA 1902

RESULT 10
LOCUS AK097650 1810 bp mRNA linear PRI 15-JUL-2002
DEFINITION Homo sapiens cDNA FLJ40331 fis, clone TESIR2031687, highly similar
to Homo sapiens mRNA for cartilage-associated protein.
ACCESSION AK097650
VERSION AK097650.1 GI:21757488
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Kanehori, K., Ishibashi, T., Chiba, Y., Fujimori, K., Hiraoka, S.,
Tanai, H., Watanabe, S., Ishida, S., Ono, Y., Hotuta, T., Watanabe, M.,
Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,
Yamamoto, J., Isono, K., Kawai-Hio, Y., Saito, K., Nishikawa, T.,
Kimura, K., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,
Wagatsuma, M., Takahashi-Fujii, A., Oshima, A., Sugiyama, A.,
Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Masuho, Y., Nagai, K.
and Isogai, T.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1810)
AUTHORS Isogai, T. and Yamamoto, J.
TITLE Direct Submission
JOURNAL Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5' - & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
FEATURES
source
1..1810
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TESIR2031687"
/issue_type="testis"
/clone_lib="TESTI2"
/note="Cloning vector: pME18SFL3"
BASE COUNT 375 a 550 c 449 g 436 t
ORIGIN
Query Match 42.8%; Score 1656; DB 9; Length 1810;
Best Local Similarity 93.1%; Pred. No. 9.9e-277;
Matches 1806; Conservative 2; Mismatches 2; Indels 130; Gaps 2;
39 CTTCTCTTCTTCCCGGCGCATGAGCCGGGGCGCGGGGGCGCGGCGCTGCTA 98
1 CTTCTCTTCTTCCCGGCGCATGAGCCGGGGCGCGGGGGCGCGGCGCTGCTA 60

```

```

99 GGGCTGCTGTGGTGGCTGGCGCGTGGCGCGCGCGCGCGCCAAATACAGAGCTACAGC 158
61 GGGCTGCTGTGGTGGCTGGCGCGTGGCGCGCGCGCGCGCCAAATACAGAGCTACAGC 120
159 TTCCGACAGTTTCCACGAGGACGAGCTGATGCGCTCGAGTCGGCTACGGGACGCGCTG 218
121 TTCCGACAGTTTCCACGAGGACGAGCTGATGCGCTCGAGTCGGCTACGGGACGCGCTG 180
219 GACAAAGTACAGCGGCGGAGCAGCTGGGCGGAGAGCSTKGGCTACCTGGAGATCAGCCTGGG 278
181 GACAAAGTACAGCGGCGGAGCAGCTGGGCGGAGAGCTGGCTACCTGGAGATCAGCCTGGG 240
279 CTGCAACGCTTGTGTCGGCGAGCAGAGGAGCTTCTCCACCGCAACTGACAGCGCGCGCG 338
241 CTGCAACGCTTGTGTCGGCGAGCAGAGGAGCTTCTCCACCGCAACTGACAGCGCGCGCG 300
339 CAGCCCGAGCGCGCGCGCTCGCAGCTATCCGAGCTGCGGCTCTTCCGGGGCGCTG 398
301 CAGCCCGAGCGCGCGCGCTCGCAGCTATCCGAGCTGCGGCTCTTCCGGGGCGCTG 360
399 CTGCGCGCGCGCAGCTGCTCAAGCGCTGCAAGCAGGCGCTGCCAGCCTTCCGCGAGTCC 458
361 CTGCGCGCGCGCAGCTGCTCAAGCGCTGCAAGCAGGCGCTGCCAGCCTTCCGCGAGTCC 420
459 CAGCCCGAGCGCGGAGTGTGCGGAGCTTCCAGCGCGCGCGCGCTTACAGTTCCTGCGAG 518
421 CAGCCCGAGCGCGGAGTGTGCGGAGCTTCCAGCGCGCGCGCGCTTACAGTTCCTGCGAG 480
519 TTGCTTTACTTCAAGGCAAAATTAATCTCCCAAGGCAATCGCGCTGCTCACACCTTTCTA 578
481 TTGCTTTACTTCAAGGCAAAATTAATCTCCCAAGGCAATCGCGCTGCTCACACCTTTCTA 540
579 CTGAACATCTGTATGACGAATATGATGAAGAGGACATGGCATATATTAAGAGCTGCTGCT 638
541 CTGAACATCTGTATGACGAATATGATGAAGAGGACATGGCATATATTAAGAGCTGCTGCT 600
639 GGTGCGGAGGACTCATTAAGACCTTGAAGAACCAAGTATGATGAAGGCTGTTCATCCGA 698
601 GGTGCGGAGGACTCATTAAGACCTTGAAGAACCAAGTATGATGAAGGCTGTTCATCCGA 659
699 GAGTGGCGGCGATACAAACGCTGAGAACTGGAGAAATCCATCACAGACATGAGAGCTGGCC 758
660 GAGTGGCGGCGATACAAACGCTGAGAACTGGAGAAATCCATCACAGACATGAGAGCTGGCC 719
759 CTTCCGAGCTTCTTCAAGGCTTTTACGAGTGTCTCGAGCGCTCGAGGTTCCAGGAG 818
720 CTTCCGAGCTTCTTCAAGGCTTTTACGAGTGTCTCGAGCGCTCGAGGTTCCAGGAG 779
819 ATCAAGGACTTCAAGGATTTTACCTTTCCATAGCAGATCATTTATGTAAGTCTTGGA 878
780 ATCAAGGACTTCAAGGATTTTACCTTTCCATAGCAG----- 816
879 TGCAAAATACAGTGTGAAGAGAACCTCACCCAGTTATAGGAGGCTATCCGTTGAGAA 938
817 ----- 816
939 TTTGTGGCTACCATGTATCATTTACTTTCAGTTTGGCTATTATAAGTTGAACACCTGAAG 998
817 -----TGAACGACCTGAAG 830
999 AATGACGCGCTGTGTCAGTACGTATCTGCTTTTGTATCAGAATGACAGGTCATGCG 1058
831 AATGACGCGCTGTGTCAGTACGTATCTGCTTTTGTATCAGAATGACAGGTCATGCG 890
1059 CAGAACCTGTGTATTACAGTACACAGGAGACCTTGGGGCTCTCGATCAGCAGCTTC 1118
891 CAGAACCTGTGTATTACAGTACACAGGAGACAGTGGGGCTCTCAGATGAGCAGCTTC 950
1119 CAGCCCGAGACCTGAAGCAGTTCAGTTCCTTTAATGTGACACACTCCAGAGGAGTGTAT 1178
951 CAGCCCGAGACCTGAAGCAGTTCAGTTCCTTTAATGTGACACACTCCAGAGGAGTGTAT 1010
1179 GACTTTGCTTAAGAAAATATATATGATGATGAGGAGAAAGTTGTGGAATATGTGGAT 1238

```


Db 1011 GACCTTCTAGGAAATATAATGGATGATGAGGAGAACTTGTGGAATATGGAT 1070
Qy 1239 GACCTTCTAGGAGAGACAGCTAGCCACAGCAACCAAGAGACTTCTCTTGG 1298
Db 1071 GACCTTCTAGGAGAGACAGCTAGCCACAGCAACCAAGAGACTTCTCTTGG 1130
Qy 1299 CGTTGAGAAACACAGATTTCTTGTCTTTTCCACAGCCAGGCTGTGATACCTCAG 1358
Db 1131 CGTTGAGAAACACAGATTTCTTGTCTTTTCCACAGCCAGGCTGTGATACCTCAG 1190
Qy 1359 AGCCTTCTTACTCTCCAAAGTGAAGGAGAGCCCTCTCTCTTAAGTGTATGCTAT 1418
Db 1191 AGCCTTCTTACTCTCCAAAGTGAAGGAGAGCCCTCTCTCTTAAGTGTATGCTAT 1250
Qy 1419 CAGGGGTGAGCGCTCTCTTCTTCACTTCACTGCGCACTGCTGATGATACCTTCTT 1478
Db 1251 CAGGGGTGAGCGCTCTCTTCTTCACTTCACTGCGCACTGCTGATGATACCTTCTT 1310
Qy 1479 CTCAACCTTTTTTTGAGATGAGTCTGCTCTCTTCCAGGCTGAGTGCATGCAAG 1538
Db 1311 CTCAACCTTTTTTTGAGATGAGTCTGCTCTCTTCCAGGCTGAGTGCATGCAAG 1370
Qy 1539 TTCTCAGTCACTCAACCTCCGCTCTTGGTTCAGCAATCTCTGTCATCAGCTCC 1598
Db 1371 TTCTCAGTCACTCAACCTCCGCTCTTGGTTCAGCAATCTCTGTCATCAGCTCC 1430
Qy 1599 CGAGTACTCTGGATTACAGGATGTCACACACCCCGGCTAATTTTGTATTTAGTAG 1658
Db 1431 CGAGTACTCTGGATTACAGGATGTCACACACCCCGGCTAATTTTGTATTTAGTAG 1490
Qy 1659 AGACGGGTTTGGATGTTGGCCAGGCTGCTGCAACTCTTCACTTCACTGATCCAT 1718
Db 1491 AGACGGGTTTGGCATGTTGGCCAGGCTGCTGCAACTCTTCACTTCACTGATCCAT 1550
Qy 1719 CTGCTTGGCTCCACAGTCTGGGATTACAGCGCTGAGCCACCATGCCCGCTCTTT 1778
Db 1551 CTGCTTGGCTCCACAGTCTGGGATTACAGCGCTGAGCCACCATGCCCGCTCTTT 1610
Qy 1779 CTCACCTTTACACCTGTTCTTATCTCTACATCTGTTTTCACACTTCATCTGCTT 1838
Db 1611 CTCACCTTTACACCTGTTCTTATCTCTACATCTGTTTTCACACTTCATCTGCTT 1670
Qy 1839 CCTCATGTCACACTGCTTCCCGATTCATAGCTGCTTCTTCACTTCACTTGGTTG 1898
Db 1671 CCTCATGTCACACTGCTTCCCGATTCATAGCTGCTTCTTCACTTCACTTGGTTG 1730
Qy 1899 AAGGCGAGTCTTCTGCTGTTGTTTTTTTGTGTTTCCAGAAATCAGTATTTTATA 1958
Db 1731 AAGGCGAGTCTTCTGCTGTTGTTTTTTGTTTCCAGAAATCAGTATTTTATA 1790
Qy 1959 AATAGAAAACATTCCTAG 1978
Db 1791 AATAGAAAACATTCCTAG 1810

RESULT 11
LOCUS BC008745
DEFINITION Homo sapiens, cartilage associated protein, clone MGC:1926
ACCESSION IMAGE:3347384, mRNA, complete cds.
VERSION BC008745.1 GI:14250580
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1439)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (25-MAY-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland,
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nri.nih.gov
Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
Lim, M., Maduro, O.L., Masiello, C., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J.,
Tiongson, E.E., Touchman, J.W., Tsurgren, C., Vogt, J.L., Walker, M.A.,
Zhang, D.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 5 Row: i Column: 24
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 3687321.

FEATURES

source

Location/Qualifiers
1..1439
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="LocusID:10491"
/db_xref="taxon:9606"
/clone="MGC:1526 IMAGE:3347384"
/tissue_type="Kidney, renal cell adenocarcinoma"
/clone_lib="NIH MGC 14"
/lab_host="DH10B-R"
/note="Vector: pOTB7"

CDS

36..1241
/codon_start=1
/product="cartilage associated protein"
/protein_id="AAH08745.1"
/db_xref="GI:14250581"
/translation="MEPGRGAAALALLCVACLAGRAQYERYSPFRBELMPL
ESVAFHDKYSGEHWASVGLTSLRHLRLSDSAFCHRNCSAAPPQFAAGLAS
YPELFLGGLARHLAKRCKGLPAPQSOPLSRDLADFORPKYFLQFAYPKANN
LPKAAAHATFELLKHPDDEMKRWAYKSLPGADYTKDLETYSLELFTRAYVYN
GENWTSITDMELAPDFKAFYECLAAECESREIKPKDFYLSIADHYVLEKIQ
CEENLTPVIGGYPEKFKVATVHYLQFAYIKLNDLKNAPCAVSILLDFQNDKVMQON
LVYYQYHRDTWGLSDEHPQPRPEAVQFENVTTLOKELYDFAKENIMDDDEGEVVEYVD
DLLELEETS"

BASE COUNT 362 a 415 c 375 g 287 t

ORIGIN

Query Match 35.4%; Score 1370.4; DB 9; Length 1439;
Best Local Similarity 99.3%; Pred. No. 2.8e-227;
Matches 1374; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 37 TTCTTCTCTCTTCTTCTGCGCGCGCGATGAGCCGCGCGCGCGCGCGCGCTGC 96
Db 10 TTCTTCTCTCTTCTTCTGCGCGCGCGATGAGCCGCGCGCGCGCGCGCTGC 69
Qy 97 TAGCGT 156
Db 70 TAGCGT 129
Qy 157 GCTTCCGAGCTTCCACGCGGACGAGCTGATCCCGCTCGAGTCGCGCTACCGGACGCGC 216
Db 130 GCTTCCGAGCTTCCACGCGGACGAGCTGATCCCGCTCGAGTCGCGCTACCGGACGCGC 189
Qy 217 TGGACAAGTACAGCGCGGAGCACTGGCGCGAGAGCACTGTGGCTACTGTGAGATCACTGCT 276
Db 190 TGGACAAGTACAGCGCGGAGCACTGGCGCGAGAGCACTGTGGCTACTGTGAGATCACTGCT 249

* 1 644: contig of 644 bp in length
* 744: gap of 100 bp
* 745
* 1428: contig of 684 bp in length
* 1429
* 1529: contig of 692 bp in length
* 2220: contig of 100 bp
* 2221: contig of 693 bp in length
* 3013: contig of 100 bp
* 3113: gap of 100 bp
* 3114: contig of 655 bp in length
* 3768: contig of 100 bp
* 3888: gap of 100 bp
* 4504: contig of 636 bp in length
* 4505: gap of 100 bp
* 5273: contig of 669 bp in length
* 5274
* 6043: gap of 100 bp
* 6143: gap of 100 bp
* 6823: contig of 680 bp in length
* 6824: gap of 100 bp
* 6924: contig of 678 bp in length
* 7601: contig of 100 bp
* 7701: gap of 100 bp
* 8380: contig of 679 bp in length
* 8381
* 8480: gap of 100 bp
* 9151: contig of 671 bp in length
* 9251: gap of 100 bp
* 9252
* 9945: contig of 694 bp in length
* 9946
* 10045: gap of 100 bp
* 10733: contig of 688 bp in length
* 10734
* 10833: gap of 100 bp
* 10834
* 11516: contig of 683 bp in length
* 11616: gap of 100 bp
* 12270: contig of 654 bp in length
* 12370: gap of 100 bp
* 13029: contig of 659 bp in length
* 13030
* 13129: gap of 100 bp
* 13796: contig of 667 bp in length
* 13896: gap of 100 bp
* 14586: contig of 690 bp in length
* 14587
* 14686: gap of 100 bp
* 15356: contig of 670 bp in length
* 15456: gap of 100 bp
* 15457
* 16108: contig of 652 bp in length
* 16109
* 16208: gap of 100 bp
* 16893: contig of 685 bp in length
* 16894
* 16933: gap of 100 bp
* 17683: contig of 690 bp in length
* 17684
* 17783: gap of 100 bp
* 18422: contig of 639 bp in length
* 18423
* 18522: gap of 100 bp
* 19212: contig of 690 bp in length
* 19213
* 19312: gap of 100 bp
* 20007: contig of 695 bp in length
* 20107: gap of 100 bp
* 20764: contig of 657 bp in length
* 20864: gap of 100 bp
* 20865
* 21523: contig of 659 bp in length
* 21524
* 21623: gap of 100 bp
* 22275: contig of 652 bp in length
* 22375: gap of 100 bp
* 23055: contig of 680 bp in length
* 23155: gap of 100 bp
* 23156
* 23803: contig of 648 bp in length
* 23804
* 23903: gap of 100 bp
* 24577: contig of 674 bp in length
* 24578
* 24579: gap of 100 bp
* 25362: contig of 685 bp in length
* 25462: gap of 100 bp
* 25463
* 26146: contig of 684 bp in length
* 26147
* 26246: gap of 100 bp
* 26247
* 26871: contig of 625 bp in length
* 26971: gap of 100 bp
* 26972
* 27661: contig of 690 bp in length
* 27662
* 27761: gap of 100 bp
* 28385: contig of 624 bp in length

* 28386
* 28486
* 29143: contig of 658 bp in length
* 29144
* 29243: gap of 100 bp
* 29244
* 29905: contig of 662 bp in length
* 29906
* 30005: gap of 100 bp
* 30682: contig of 677 bp in length
* 30683
* 30782: gap of 100 bp
* 30783
* 31455: contig of 673 bp in length
* 31456
* 31555: gap of 100 bp
* 32227: contig of 672 bp in length
* 32228
* 32327: gap of 100 bp
* 33011: contig of 684 bp in length
* 33012
* 33111: gap of 100 bp
* 33800: contig of 689 bp in length
* 33812
* 33900: gap of 100 bp
* 33901
* 34590: contig of 690 bp in length
* 34591
* 34690: gap of 100 bp
* 35378: contig of 688 bp in length
* 35379
* 35478: gap of 100 bp
* 36135: contig of 657 bp in length
* 36136
* 36235: gap of 100 bp
* 36236
* 36893: contig of 658 bp in length
* 36894
* 36993: gap of 100 bp
* 37665: contig of 672 bp in length
* 37666
* 37765: gap of 100 bp
* 38433: contig of 688 bp in length
* 38434
* 38533: gap of 100 bp
* 38534
* 39217: contig of 684 bp in length
* 39218
* 39317: gap of 100 bp
* 39318
* 39996: contig of 679 bp in length
* 39997
* 40097
* 40787
* 40886: gap of 100 bp
* 41547: contig of 661 bp in length
* 41548
* 41648: gap of 100 bp
* 42337: contig of 690 bp in length
* 42338
* 42437: gap of 100 bp
* 42438
* 43120
* 43129: gap of 100 bp
* 43913: contig of 694 bp in length
* 43914
* 44013: gap of 100 bp
* 44014
* 44660: contig of 647 bp in length
* 44661
* 44760: gap of 100 bp
* 44761
* 45430: contig of 670 bp in length
* 45431
* 45530: gap of 100 bp
* 45531
* 46207: contig of 677 bp in length
* 46208
* 46307: gap of 100 bp
* 46308
* 46986: contig of 679 bp in length
* 46987
* 47086: gap of 100 bp
* 47087
* 47668: contig of 682 bp in length
* 47669
* 47769: gap of 100 bp
* 47869
* 48226: contig of 658 bp in length
* 48227
* 48627: gap of 100 bp
* 48628
* 49307: contig of 680 bp in length
* 49308
* 49406: gap of 100 bp
* 49407
* 50055: contig of 649 bp in length
* 50056
* 50155: gap of 100 bp
* 50156
* 50845: contig of 690 bp in length
* 50846
* 50945: gap of 100 bp
* 50946
* 51649
* 51749
* 52449: contig of 701 bp in length
* 52450
* 52549: gap of 100 bp
* 52550
* 53209: contig of 660 bp in length
* 53210
* 53309: gap of 100 bp
* 53310
* 53946: contig of 637 bp in length
* 53947
* 54046: gap of 100 bp
* 54047
* 54704: contig of 658 bp in length

Query Match 13.2%; Score 510.2; DB 2; Length 64706;
Best Local Similarity 99.4%; Pred. No. 2.2e-78;
Matches 520; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2004, 07:10:40 ; Search time 4130.11 Seconds
(without alignments)
11232.506 Million cell updates/sec

Title: US-09-729-674-1_COPY_132_1265

Perfect score: 1134

Sequence: 1 gggcgcccaatacgaacg.....tgggaactggaggagaccagc 1134

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.:

1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_cm.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rnd.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1133.2	99.9	2307	9 HSAJ6470	AJ006470 Homo sapi
2	1133.2	99.9	3871	6 BD140451	BD140451 Secreted
3	1128.4	99.5	1439	9 BC008745	BC008745 Homo sapi
4	1128.4	99.5	1961	6 BD156517	BD156517 Primer fo
5	1128.4	99.5	1961	6 AK001634	AK001634 Homo sapi
6	1118	98.6	2010	9 AK075401	AK075401 Homo sapi
7	1042.4	91.9	1922	9 AK091772	AK091772 Homo sapi
8	1019.6	89.9	1903	9 AK092990	AK092990 Homo sapi
9	897.2	79.1	1665	10 BC049890	BC049890 Mus muscu
10	895.6	79.0	1685	10 MNAJ6469	AJ006469 Mus muscu
11	850	75.0	1810	9 AK097650	AK097650 Homo sapi
12	587.6	51.8	818	6 BD148800	BD148800 Primer fo
13	494.6	43.6	1665	5 GGCARP	X97607 G.gallus mr
14	455.4	40.2	1745	9 BC011701	BC011701 Homo sapi
15	455.4	40.2	1889	9 BC001047	BC001047 Homo sapi
16	455.4	40.2	2037	9 BC007942	BC007942 Homo sapi
17	455.4	40.2	2347	6 AX335490	AX335490 Sequence
18	455.4	40.2	2347	6 HSU47621	U47621 Homo sapien
19	455.4	40.2	2443	6 AX714040	AX714040 Sequence
20	455.4	40.2	2443	9 AK056085	AK056085 Homo sapi
21	453.8	40.0	2079	9 HSA250583	AJ250583 Homo sapi
22	418.2	36.9	1903	10 BC031856	BC031856 Mus muscu
23	406.6	35.9	182509	9 AC112211	AC112211 Homo sapi
24	403	35.5	1407	10 RNSG65MR	X55454 R.norvegicu
25	379.2	33.4	64706	2 AC136358	AC136358 Homo sapi
26	320.8	28.3	226615	2 AC122652	AC122652 Rattus no
27	193.6	17.1	178345	9 AC091172	AC091172 Homo sapi
28	193.6	17.1	196735	2 AC012192	AC012192 Homo sapi
29	169.2	14.9	176789	2 AC032013	AC032013 Mus muscu
30	169.2	14.9	217092	10 AL590968	AL590968 Mouse DNA
31	166	14.6	226822	2 AC112575	AC112575 Rattus no
32	166	14.6	241607	2 AC113636	AC113636 Rattus no
33	156.4	13.8	2524	6 BD156505	BD156505 Primer fo
34	156.4	13.8	2524	9 AK027680	AK027680 Homo sapi
35	156.4	13.8	2563	6 AX136191	AX136191 Sequence
36	156.4	13.8	2563	6 BD123548	BD123548 Secretary
37	156.4	13.8	2563	9 AK075418	AK075418 Homo sapi
38	156.4	13.8	2583	6 AX704765	AX704765 Sequence
39	156.4	13.8	2585	6 BD156581	BD156581 Primer fo
40	156.4	13.8	2585	9 AK027697	AK027697 Homo sapi
41	156.4	13.8	2993	6 BD158280	BD158280 Primer fo
42	156.4	13.8	2993	9 AK027648	AK027648 Homo sapi
43	154.8	13.7	2600	9 AF097432	AF097432 Homo sapi
44	154.8	13.7	2829	9 AF097431	AF097431 Homo sapi
45	147.4	13.0	2249	10 NMU430350	AJ430350 Mus muscu

ALIGNMENTS

RESULT 1
HSAJ6470

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

HSAJ6470

Homo sapiens mRNA for cartilage-associated protein (CASP).

AJ006470

GI:3687321

cartilage-associated protein; CASP.

Homo sapiens (human)

Homo sapiens

2307 bp

linear

PRI 23-MAR-2000

REFERENCES

1

Tonachini L., Morello R., Monticone M., Skaug J., Scherer S.W.,

Cancedda R. and Castagnola P.

CDNA cloning, characterization and chromosome mapping of the gene

```
JOURNAL encoding human cartilage associated protein (CRTAP)
MEDLINE Cytogenet. Cell Genet. 87 (3-4), 191-194 (1999)
20169181
PUBMED 10702564
REFERENCE 2 (bases 1 to 2307)
AUTHORS Castagnola, P.
DIRECT SUBMISSION
TITLE Submitted (12-JUN-1998) Castagnola P., Biotechnology in Oncologia,
JOURNAL Ist. Naz.le per la Ricerca sul Cancro, L.go R. Benzi 10 Genoa,
16132, ITALY
FEATURES
    source Location/Qualifiers
        1..2307
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="Contig of pCPH2 and I.M.A.G.E. cloneID 1090695"
        /tissue_type="fetal brain"
        /dev_stage="19-23 weeks-old"
    gene 1..2307
        /gene="casp"
    CDS 12..1217
        /gene="casp"
        /codon_start=1
        /product="cartilage-associated protein (CASP)"
        /protein_id="CAA07054.1"
        /db_xref="GI:3687322"
        /translation="MEFGRGAAALALLVCALRAGRAQYERYFRSPPRDELMLPL
        ESAYRHADKYSGHVESYGLIEISLRHLRLDSFAFCHRNCSAOPPEPAAGLAS
        YPELRFGGLRLAKHPDCKOGKLPQSPQREVLADFORPEPVKFLQFAVFKANN
        LPKAIAAHFTLLKHPDCKOGKLPQSPQREVLADFORPEPVKFLQFAVFKANN
        GENWRTSITDMELALPDPFKAFVECLAACGSRREIKDPDFYLSIADHVVELEKIQ
        CEENLTPVIGGFVPEKFAVTMIHYLQAFYKLNLDLKNAPCAVSYLFDONDKWMQCN
        LVYYQHRDWTGSLDEHFQRPPEAVQFNFTLQKELYDFAKENIMDDDEGEVVEYVD
        DLELEETS"
    sig_peptide 12..89
    mat_peptide 90..1214
        /gene="casp"
        /product="cartilage-associated protein (CASP)"
BASE COUNT 545 a 633 c 548 g 578 t 3 others
ORIGIN
Query Match 99.9%; Score 1133.2; DB 9; Length 2307;
Best Local Similarity 99.8%; Pred No. 2.4e-180;
Matches 1132; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGCGCGCCCAATACGAACGCTACAGCTTCGCGAGCTTCCACGGGACGAGCTGATGCGG 60
Db 81 GGGCGCGCCCAATACGAACGCTACAGCTTCGCGAGCTTCCACGGGACGAGCTGATGCGG 140
QY 61 CTCGAGTCGGCTACCGGACGCGCTGGACAGTACAGCGGCGGAGCACTGGCGCGAGAGC 120
Db 141 CTCGAGTCGGCTACCGGACGCGCTGGACAGTACAGCGGCGGAGCACTGGCGCGAGAGC 200
QY 121 STKGCTACCTGGAGATCAGCTGCGGCTGCACCGCTTGTGCGCGACGAGCGGCGCTTC 180
Db 201 GTGGGCTACCTGGAGATCAGCTGCGGCTGCACCGCTTGTGCGCGACGAGCGGCGCTTC 260
QY 181 TGCACCGCAACTGACGCGCGCGCGGACGCGGCGGAGCGGCGGCGGCGGCTGCGAGCTAT 240
Db 261 TGCACCGCAACTGACGCGCGCGCGGAGCGGCGGAGCGGCGGCGGCGGCTGCGAGCTAT 320
QY 241 CCGGAGCTGCGGCTCTTGGGGGCGCTGTGCGCGCGGCGGAGCGGCGGCTCAAGCGTGAAG 300
Db 321 CCGGAGCTGCGGCTCTTGGGGGCGCTGTGCGCGCGGCGGAGCGGCGGCTGCGAGCTGCAAG 380
QY 301 CAGGCGCTGCGAGCTTCCGCGAGTCCAGCGGCGGAGCGGCGGAGCTGCGGAGCTTCCAG 360
Db 381 CAGGCGCTGCGAGCTTCCGCGAGTCCAGCGGCGGAGCGGCGGAGCTGCGGAGCTTCCAG 440
QY 361 CCGCGCGGCGGCGGCTTCCGCGAGTCCGAGTTCGCTTCAAGGCAAAATATCTCCGCCAAA 420
Db 441 CCGCGCGGCGGCGGCTTCCGCGAGTTCGCTTCAAGGCAAAATATCTCCGCCAAA 500
QY 421 GCATTCGCGGCTGTCTCACACCTTTCTACTGAAGCATCTGATGACGAAATGATCAAGAGG 480
Db 501 GCATTCGCGGCTGTCTCACACCTTTCTACTGAAGCATCTGATGACGAAATGATCAAGAGG 560
QY 481 AACATGGCATATTAAGAGCGCTGCTGGTCCGAGGACTACATTAAGAGCTGGAACCC 540
Db 561 AACATGGCATATTAAGAGCGCTGCTGGTCCGAGGACTACATTAAGAGCTGGAACCC 620
QY 541 AAGTCATATGAAGCGCTGCTTCATCCGAGCAGTGGGGCATACCAACGGTGAGAACTGGAGA 600
Db 621 AAGTCATATGAAGCGCTGCTTCATCCGAGCAGTGGGGCATACCAACGGTGAGAACTGGAGA 680
QY 601 ACATCCATCACAGCATGAGCTGCGCCTTCCGAGCTTCTTCAAGGCTTTTACAGTGT 660
Db 681 ACATCCATCACAGCATGAGCTGCGCCTTCCGAGCTTCTTCAAGGCTTTTACAGTGT 740
QY 661 CTCGAGCGCTCGAGGGTTCAGGAGGATCAAGGACTCAAGGATTTCTACCTTTCCATA 720
Db 741 CTCGAGCGCTCGAGGGTTCAGGAGGATCAAGGACTTTCAAGGATTTCTACCTTTCCATA 800
QY 721 GCAGATCATATTGTAGAAAGTTCTGGAATGCAAAATACAGTGTGAAGAGAACTCACCCCA 780
Db 801 GCAGATCATATTGTAGAAAGTTCTGGAATGCAAAATACAGTGTGAAGAGAACTCACCCCA 860
QY 781 GTTATAGAGGCTATCCGGTTGAGAAATTTGCGCTACCATGATCATTTACCTTCAGTGT 840
Db 861 GTTATAGAGGCTATCCGGTTGAGAAATTTGCGCTACCATGATCATTTACCTTCAGTGT 920
QY 841 GCCTATTATAGTTCAACGAGCTGGAAGAACTGAGCCCTGTCAGTGTGAGTATCTGCTC 900
Db 921 GCCTATTATAGTTCAACGAGCTGGAAGAACTGAGCCCTGTCAGTGTGAGTATCTGCTC 980
QY 901 TTGTATCAGATGACAAGGTCATGACGAGAACTGCTGTGTGTTATACCATGACGAGGAC 960
Db 981 TTGTATCAGATGACAAGGTCATGACGAGAACTGCTGTGTGTTATACCATGACGAGGAC 1040
QY 961 ACTTGGGCGCTCTCGGATGAGCACTTCCAGCCGAGCTGAGGAGTTCAGTTCCTTTAAT 1020
Db 1041 ACTTGGGCGCTCTCGGATGAGCACTTCCAGCCGAGCTGAGGAGTTCAGTTCCTTTAAT 1100
QY 1021 GTGACCACTCCAGAGGAGCTGTATGACTTTGTAAGGAAATATATATGATGATGAT 1080
Db 1101 GTGACCACTCCAGAGGAGCTGTATGACTTTGTAAGGAAATATATATGATGATGAT 1160
QY 1081 GAGGAGAGGTTGTGGAATATGTGATGACCTCTTGGAACTGGAGGAGACGAGC 1134
Db 1161 GAGGAGAGGTTGTGGAATATGTGATGACCTCTTGGAACTGGAGGAGACGAGC 1214
RESULT 2
BD140451 3871 bp DNA linear PAT 18-SEP-2002
LOCUS Secreted proteins and polynucleotides encoding them.
DEFINITION BD140451
ACCESSION BD140451.1 GI:23235396
VERSION JP 2002506611-A/1.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3871)
AUTHORS Jacobs,K., McCoy,J.M., Lavallie,E.R., Racie,L.A.C., Evans,C.,
Merberg,D., Treacy,M., Agostino,M.J., II,R.J.S., Wong,G.G.,
Clark,H.F. and Fechtel,K.
TITLE Secreted proteins and polynucleotides encoding them
JOURNAL Patent: JP 2002506611-A 1 05-MAR-2002,
GENETICS INSTITUTE INC
COMMENT CS Homo sapiens (human)
PN JP 2002506611-A/1
PD 05-MAR-2002
PF 24-NOV-1998 JP 2000522118
```


PR 26-NOV-1997 US 60/066804,23-NOV-1998 US 09/197886 PI
KENNETH JACOBS, JOHN M MCCOY, EDWARD R LAVALLIE, LISA A COLLINS PI
PI CHERYL EVANS, DAVID MERBERG, MAURICE TREACY, MICHAEL J AGOSTINO,
PI ROBERT J STEININGER II, GORDON G WONG, HILARY F CLARK, KIM PI
FECHTEL
PC C12N15/09, C07K14/00, C12N1/21, C12N5/10, C12P19/34, C12P21/02, PC
C12Q1/68//
PC A61P29/00, A61P35/00, A61P37/04, A61P37/06, C12N15/00, C12N5/00 CC
Secreted proteins and polynucleotides encoding them FH Key
Location/Qualifiers
FT source 1..3871
FT Location/Qualifiers
1..3871
/organism="Homo sapiens (human)"
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 965 a 913 c 1033 g 958 t 2 others
ORIGIN
Query Match 99.9%; Score 1133.2; DB 6; Length 3871;
Best Local Similarity 100.0%; Pred. No. 2.4e-180;
Matches 1134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGGCGCGCCCAATACGAACGCTACAGCTTCGGGAGCTTCCACGGGACGAGCTGATGCCG 60
Db 132 GGGCGCGCCCAATACGAACGCTACAGCTTCGGGAGCTTCCACGGGACGAGCTGATGCCG 191
Qy 61 CTCGAGTCGGCTACCGCGACGCGCTGACAAAGTACAGCGCGGAGCACTGGGCGGAGAGC 120
Db 192 CTCGAGTCGGCTACCGCGACGCGCTGACAAAGTACAGCGCGGAGCACTGGGCGGAGAGC 251
Qy 121 STKGCTACCTGGAGATCAGCTTCGGGCTCGACCGCTTGTGCGCGACGACGAGGCTTC 190
Db 252 STKGCTACCTGGAGATCAGCTTCGGGCTCGACCGCTTGTGCGCGACGACGAGGCTTC 311
Qy 181 TGCCACCGCACTGACGCGCGCGCGCGAGCCGCGCGCGCGCGCGCGCGCGCGCGCTAT 240
Db 312 TGCCACCGCACTGACGCGCGCGCGCGAGCCGCGCGCGCGCGCGCGCGCGCGCGCTAT 371
Qy 241 CCCGAGCTGGCGCTCTTCGGGGGCGCTGCTGCGCGCGCGCGCACTGCTCAAGCGCTGCAAG 300
Db 372 CCCGAGCTGGCGCTCTTCGGGGGCGCTGCTGCGCGCGCGCGCACTGCTCAAGCGCTGCAAG 431
Qy 301 CAGGCGCTGCGAGCTTCGCGAGTCCGCGAGTCCGCGAGTCCGCGAGTGTGCGCGAGCTTCAG 360
Db 432 CAGGCGCTGCGAGCTTCGCGAGTCCGCGAGTCCGCGAGTGTGCGCGAGCTTCAG 491
Qy 361 CGCGCGAGCGCTACAAAGTTCCTGAGTTCGCTTCAAGGCAATATCTCCGCAAA 420
Db 492 CGCGCGAGCGCTACAAAGTTCCTGAGTTCGCTTCAAGGCAATATCTCCGCAAA 551
Qy 421 GCCATCGCGCTGCTCAACCTTTCTACTGAAGCATCTCTGATGACGAAATGATGAAGAGG 480
Db 552 GCCATCGCGCTGCTCAACCTTTCTACTGAAGCATCTCTGATGACGAAATGATGAAGAGG 611
Qy 481 AACATGCATATTATAGAGCTGCTGCTGCGGAGCTACATTTAAGACCTGGAAACC 540
Db 612 AACATGCATATTATAGAGCTGCTGCTGCGGAGCTACATTTAAGACCTGGAAACC 671
Qy 541 AAGTCATATGAAGCGCTGTTTCATCCGAGCAGTGGCGGCATACACCGTGAAGACTGGAGA 600
Db 672 AAGTCATATGAAGCGCTGTTTCATCCGAGCAGTGGCGGCATACACCGTGAAGACTGGAGA 731
Qy 601 ACATCCATCAGACATGAGCTGGCGCTTCCGAGCTTCTTCAAGCGCTTTTACGAGTGT 660
Db 732 ACATCCATCAGACATGAGCTGGCGCTTCCGAGCTTCTTCAAGCGCTTTTACGAGTGT 791
Qy 661 CTCGAGCGCTCGAGGCTTCCAGGAGATCAAGGACTTCAAGGATTTCTACCTTTCCATA 720
Db 792 CTCGAGCGCTCGAGGCTTCCAGGAGATCAAGGACTTCAAGGATTTCTACCTTTCCATA 851

QY 721 GCAGATCATATTGTAGAAAGTTCTTGGATGCAAAAATACAGTGTGAAGAGAACTCACCCCA 780
DB 852 GCAGATCATATTGTAGAAAGTTCTTGGATGCAAAAATACAGTGTGAAGAGAACTCACCCCA 911
QY 781 GTTATAGAGGCTATCCCGTTGAGAAATTTGGCTACCATGTATCATCTTGCAGTTT 840
DB 912 GTTATAGAGGCTATCCCGTTGAGAAATTTGGCTACCATGTATCATCTTGCAGTTT 971
QY 841 GCCTATTATTAAGTTGAAACGACTGAAAGTAATGAGCCCTGTGCACTGAGTATCTGCTC 900
DB 972 GCCTATTATTAAGTTGAAACGACTGAAAGTAATGAGCCCTGTGCACTGAGTATCTGCTC 1031
QY 901 TTTGATCAGATGACAAGGTCATGACGAGCAACCTGTGTGTATTACCATACACAGGAC 960
DB 1032 TTTGATCAGATGACAAGGTCATGACGAGCAACCTGTGTGTATTACCATACACAGGAC 1091
QY 961 ACTTGGGCGCTCTCGGATGAGCACTTCCAGCCGAGCTGAGGAGTTCAGTTCTTTTAA 1020
DB 1092 ACTTGGGCGCTCTCGGATGAGCACTTCCAGCCGAGCTGAGGAGTTCAGTTCTTTTAA 1151
QY 1021 GTGACCACACTCCAGAAGGAGCTGTATGACTTTGCTAAGGAAATATATATGATGATGAT 1080
DB 1152 GTGACCACACTCCAGAAGGAGCTGTATGACTTTGCTAAGGAAATATATATGATGATGAT 1211
QY 1081 GAGGAGAGAGTTGTGGAATATGTGATGACCTCTTGGAACTGGAGGAGACCGC 1134
DB 1212 GAGGAGAGAGTTGTGGAATATGTGATGACCTCTTGGAACTGGAGGAGACCGC 1265
RESULT 3
BC008745
LOCUS
DEFINITION Homo sapiens, cartilage associated protein, clone MGC:1926
IMAGE:3347384, mRNA, complete cds.
ACCESSION BC008745
VERSION BC008745.1 GI:14250580
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1439)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (25-MAY-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/PTP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@hgrl.nih.gov
Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
Lim, M., Maduro, Q.L., Masiello, C., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Snyder, B., Stantropop, S., Thomas, P.J.,
Tingerson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,
Zhang, J.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IMAGE Plate: 5 Row: 1 Column: 24
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 3687321.
Location/Qualifiers

```
source
1. .1439
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="LOCUS:10491"
/db_xref="taxon:9606"
/clone="MGC:1926 IMAGE:3347384"
/tissue_type="kidney, renal cell adenocarcinoma"
/clone_lib="NH_MGC_14"
/lab_host="DH10B-R"
/notes="vector: pOTB7"
36. .1241
/codon_start=1
/product="cartilage associated protein"
/protein_id="AA08745.1"
/db_xref="GI:14250581"
/translation="MPEGRGCAALLALLVACALRAGRAQVRSFRSPRDELMLPL
ESAYRALDKYSGEHWAESVGLYLSRLRLHLDSEAFCHNCRAPOPEPAAGLAS
YPERLFGLLRLRRKCKGKGFAPRQPSQDLVADQREPKYKFLQFAYFRANN
LPKAAAHATFLKHPDDEMMKNMAYIKLPGAEYIKOLETKSYESULFIRAVRAN
GENWRTSITMELALPDDFKAFYECALACGSEIKDFKDFYLSIADHYVELECKIQ
CEENLVITGGYPPVEKFTVMHYLQFAYVYKLDKNAAPCAVSYLLFDQNDKVMQON
LVYQYHRDWTGLSDEHFQPRPEAVQFNVTTLKELYDPAKENIMDDDEGEVVEVD
DLLELETS"
BASE COUNT 362 a 415 c 375 g 287 t
ORIGIN
Query Match 99.5%; Score 1128.4; DB 9; Length 1439;
Best Local Similarity 99.6%; Pred. No. 1.6e-179;
Matches 1129; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 GGGCGCGCCCAATACGAACGTACAGCTTCGCGAGCTTCCACGGGACGAGCTGATGCCG 60
DB 105 GGGCGCGCCCAATACGAACGTACAGCTTCGCGAGCTTCCACGGGACGAGCTGATGCCG 164
QY 61 CTCGAGTCGGCTACCGGACGCGCTGCACAAGTACAGCGCGGACGACTGGCGCGAGAGC 120
DB 165 CTCGAGTCGGCTACCGGACGCGCTGCACAAGTACAGCGCGGACGACTGGCGCGAGAGC 224
QY 121 STKGCTACCTGGAGATCAGCTTCGGGCTGCACCGCTTGTGCGCGACGAGGCGCTTC 180
DB 225 GTGGGCTACCTGGAGATCAGCTTCGGGCTGCACCGCTTGTGCGCGACGAGGCGCTTC 284
QY 181 TGCCACCGCAACTGACGCGCGCGCGCGAGCCGCGCGCGCGCTGCGCGAGCTAT 240
DB 285 TGCCACCGCAACTGACGCGCGCGCGCGAGCCGCGCGCGCGCTGCGCGAGCTAT 344
QY 241 CCGGAGCTGCGCTTCTCGGGGCGCTGTGCGCGCGCGCGCTGCGCGAGCTGCGCGAG 300
DB 345 CCGGAGCTGCGCTTCTCGGGGCGCTGTGCGCGCGCGCGCTGCGCGAGCTGCGCGAG 404
QY 301 CAGGCGCTGCGAGCTTCCGCGAGTCCCGAGCCGCGCGCGCGAGGCTGCGCGAGCTCCAG 360
DB 405 CAGGCGCTGCGAGCTTCCGCGAGTCCCGAGCCGCGCGCGAGGCTGCGCGAGCTCCAG 464
QY 361 CGCGCGAGCGCTTACAGTTCCTGAGTTCTGCTTACTTCAAGGCAAAATATCTCCCAAA 420
DB 465 CGCGCGAGCGCTTACAGTTCCTGAGTTCTGCTTACTTCAAGGCAAAATATCTCCCAAA 524
QY 421 GCATCGCGGCTGCTACACCTTCTGAGTTCCTGAGTTCCTGAGTTCCTGAGTTCCTGAG 480
DB 525 GCATCGCGGCTGCTACACCTTCTGAGTTCCTGAGTTCCTGAGTTCCTGAGTTCCTGAG 584
QY 481 AACATGGCATATTATAAGAGCTGCTGCTGCGCGAGGACTACATTAAGAGCTGGAAC 540
DB 585 AACATGGCATATTATAAGAGCTGCTGCTGCGCGAGGACTACATTAAGAGCTGGAAC 644
QY 541 AAGTCATATGAAGCTGTTTCATCGAGAGCTGCGGCGATACACCGTGTGAAGCTGGA 600
DB 645 AAGTCATATGAAGCTGTTTCATCGAGAGCTGCGGCGATACACCGTGTGAAGCTGGA 704
QY 601 ACATCCATCAGAGATGAGGCTGCGCTTCCGACTTCTTCAAGCGCTTTTACGAGTGT 660
DB 705 ACATCCATCAGAGATGAGGCTGCGCTTCCGACTTCTTCAAGCGCTTTTACGAGTGT 764
```

```
661 CTCGAGCCTCGAGGGTTCCAGGAGATCAAGACTTCAAGGATTTCTACCTTCCATA 720
DB 765 CTCGAGCCTCGAGGGTTCCAGGAGATCAAGGATTTCTACCTTCCATA 824
QY 721 GCAGATCATATTATGATAGAGTTCTGGAATGGAATAATACAGTGTGAAGAGAACTACACCCA 780
DB 825 GCAGATCATATTATGATAGAGTTCTGGAATGGAATAATACAGTGTGAAGAGAACTACACCCA 884
QY 781 GTTATAGGAGGCTATCCGGTTGAGAAATTTGCTGGTACCATGATCATTTAGTTCAGTTT 840
DB 885 GTTATAGGAGGCTATCCGGTTGAGAAATTTGCTGGTACCATGATCATTTAGTTCAGTTT 944
QY 841 GCTATTATTAAGTTGAACGAGCTGAAGAAATCAGCCCTGTGAGTCAAGTATCTGTC 900
DB 945 GCTATTATTAAGTTGAACGAGCTGAAGAAATCAGCCCTGTGAGTCAAGTATCTGTC 1004
QY 901 TTTGATCAGATGACAAAGTTCATGACGAGAACTGCTGTATTTACCATGACACAGGAC 960
DB 1005 TTTGATCAGATGACAAAGTTCATGACGAGAACTGCTGTATTTACCATGACACAGGAC 1064
QY 961 ACTTGGGCGCTCTCGGATGAGCACTTCCAGCCCGAGCTTGGAACTGGAGGAGACCA 1020
DB 1065 ACTTGGGCGCTCTCGGATGAGCACTTCCAGCCCGAGCTTGGAACTGGAGGAGACCA 1124
QY 1021 GTGACACACTCCAGAGAGCTGTATGACTTGTCTAAGGAAATATATGATGATGAT 1080
DB 1125 GTGACACACTCCAGAGAGCTGTATGACTTGTCTAAGGAAATATATGATGATGAT 1184
QY 1081 GAGGAGGAAGTTGTGGAATATGTGATGACCTTCTTGGAACTGGAGGAGACCA 1134
DB 1185 GAGGAGGAAGTTGTGGAATATGTGATGACCTTCTTGGAACTGGAGGAGACCA 1238
```

```
RESULT 4
BD156517 1961 bp DNA linear PAT 17-JAN-2003
LOCUS Primer for synthesizing full-length cDNA and use thereof.
DEFINITION BD156517
ACCESSION BD156517
VERSION BD156517.1 GI:27862275
KEYWORDS JP 2002191363-A/11360.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1961)
Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
Primer for synthesizing full-length cDNA and use thereof.
Patent: JP 2002191363-A 11360 09-JUL-2002;
HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002191363-A/11360
PD 09-JUL-2002
PF 28-JUL-2000 JP 2000280990
PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
PI SAITO,
PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
PI KEIICHI NAGAI,TETSUJI OTSUKI
PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
10,
PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof PH Key
Location/Qualifiers
(40) .. (1242) .
FT CDS Location/Qualifiers
1. .1961
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 417 a 579 c 481 g 484 t
ORIGIN
```

Query Match		99.5%;	Score 1128.4;	DB 6;	Length 1961;
Best Local Similarity		99.6%;	Pred. No. 1.6e-179;		
Matches 1129;		Conservative	2;	Mismatches	3; Indels 0; Gaps 0;
QY	1	GGGCGCGCCCAATACAAAGCTTACAGCTTCCGAGCTTCCACGGGACGAGCTGATGCGG	60		
DB	109	GGGCGCGCCCAATACAAAGCTTACAGCTTCCGAGCTTCCACGGGACGAGCTGATGCGG	168		
QY	61	CTCGAGTCCGCTTACCGGACCGCTGGACAAGTACAGCGGCGAGCACTGGGCCGAGAGC	120		
DB	169	CTCGAGTCCGCTTACCGGACCGCTGGACAAGTACAGCGGCGAGCACTGGGCCGAGAGC	228		
QY	121	STKGGTACTCGAGATCAGCTTCCGAGCTTCCACGGGACGAGCTTCCACGGGACGAGC	180		
DB	229	GTGGGCTACTCGAGATCAGCTTCCGAGCTTCCACGGGACGAGCTTCCACGGGACGAGC	288		
QY	181	TGCCACCGCAACTGACGAGCGCGCGAGCGGCGAGCGCGCGCGCGCGCGCGCGCGCG	240		
DB	289	TGCCACCGCAACTGACGAGCGCGCGAGCGGCGAGCGCGCGCGCGCGCGCGCGCGCG	348		
QY	241	CCCGAGCTCGGCTTCTTCGCGGCGCTGCTGCGCGCGCGCACTGCTCAAGCGCTGCAAG	300		
DB	349	CCCGAGCTCGGCTTCTTCGCGGCGCTGCTGCGCGCGCGCACTGCTCAAGCGCTGCAAG	408		
QY	301	CAGGCGCTGCGAGCTTCCGCGAGTCCGAGCGCGGCGAGGCTGCTGCGGAGCTTCCAG	360		
DB	409	CAGGCGCTGCGAGCTTCCGCGAGTCCGAGCGCGGCGAGGCTGCTGCGGAGCTTCCAG	468		
QY	361	CGCGCGGAGCGCTTACAGTCTTCTGCGAGTTCGCTTCTCAAGGCAATATCTCCCAAA	420		
DB	469	CGCGCGGAGCGCTTACAGTCTTCTGCGAGTTCGCTTCTCAAGGCAATATCTCCCAAA	528		
QY	421	GCATCGCGCTGCTCAGACCTTCTTCTGAGCATCTGATGACGAAATGATGAGAGG	480		
DB	529	GCCATCGCGCTGCTCAGACCTTCTTCTGAGCATCTGATGACGAAATGATGAGAGG	588		
QY	481	ACATCGCATATTAAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	540		
DB	589	ACATCGCATATTAAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	648		
QY	541	AAGTCATATGAAGCTGTTTCATCCGAGAGTCCGCGGATACAAACGGTGAGAACTGGAGA	600		
DB	649	AAGTCATATGAAGCTGTTTCATCCGAGAGTCCGCGGATACAAACGGTGAGAACTGGAGA	708		
QY	601	ACATCATACAGACATGAGCTTCCGCGCTTCCGAGCTTCTTCAAGCTTTTACAGTGT	660		
DB	709	ACATCATACAGACATGAGCTTCCGCGCTTCCGAGCTTCTTCAAGCTTTTACAGTGT	768		
QY	661	CTCGAGCTCGGAGGTTCCAGGAGATCAAGGACTTCAAGGATTTCTACCTTTCCATA	720		
DB	769	CTCGAGCTCGGAGGTTCCAGGAGATCAAGGACTTCAAGGATTTCTACCTTTCCATA	828		
QY	721	GCAGATCATTTAGTGTGGAATGCAAAATACAGTGTGAAGAACTTCAACCCA	780		
DB	829	GCAGATCATTTAGTGTGGAATGCAAAATACAGTGTGAAGAACTTCAACCCA	888		
QY	781	GTATAGGAGCTTACCGTTGAGAAATTTGGCTACCATGATCATTTACTTGCAGTTT	840		
DB	889	GTATAGGAGCTTACCGTTGAGAAATTTGGCTACCATGATCATTTACTTGCAGTTT	948		
QY	841	GCCTATTATTAAGTTGAACGACCTGAAGAAATGCAAGCTTCCGAGCTTCTGCTC	900		
DB	949	GCCTATTATTAAGTTGAACGACCTGAAGAAATGCAAGCTTCCGAGCTTCTGCTC	1008		
QY	901	TTTGATCAGATGACAAAGTTCAGCAGAACTTGGTGTATTACAGTACCAAGGAC	960		
DB	1009	TTTGATCAGATGACAAAGTTCAGCAGAACTTGGTGTATTACAGTACCAAGGAC	1068		
QY	961	ACTTGGGCGCTCTCGGATGAGCACTTCCAGCCAGACCTGAAGCAAGTTCAGTCTTTAAT	1020		
DB	1069	ACGTGGGCGCTCTCAGATGAGCACTTCCAGCCAGACCTGAAGCAAGTTCAGTCTTTAAT	1128		
QY	1021	GTGACCACTCCAGAGGAGCTGTATGACTTGTCTAAGGAAATATATATGATGATGAT	1080		
DB	1129	GTGACCACTCCAGAGGAGCTGTATGACTTGTCTAAGGAAATATATATGATGATGAT	1188		
QY	1081	GAGGAGAACTTGGTAATATGCGTACCTTGGAACTGGAGGAGGAGC	1134		
DB	1189	GAGGAGAACTTGGTAATATGCGTACCTTGGAACTGGAGGAGGAGC	1242		
RESULT 5					
AK001634	1961 bp	mRNA	linear	PRI 01-AUG-2002	
LOCUS					
DEFINITION					
Homo sapiens cDNA FLJ10772 fis, clone NT2RP4000243, highly similar					
to Homic sapiens mRNA for cartilage-associated protein.					
ACCESSION					
AK001634					
VERSION					
AK001634.1 GI:7023007					
KEYWORDS					
oligo capping; fis (full insert sequence).					
SOURCE					
Homo sapiens (human)					
ORGANISM					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
REFERENCE					
1					
Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y.,					
Nishikawa, T., Nagai, K., Sugano, S., Aotaka, S., Yoshikawa, Y.,					
Matsumura, H., Tehii, S., Kawai, Y., Saito, K., Yamamoto, J.,					
Wakamatsu, A., Nakamura, Y., Nagahari, K., Masuko, Y. and Sasaki, N.					
NEDO human cDNA sequencing project					
Unpublished					
2 (bases 1 to 1961)					
Isogai, T. and Otsuki, T.					
Direct Submission					
Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute,					
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan					
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)					
NEDO human cDNA sequencing project supported by Ministry of					
International Trade and Industry of Japan; cDNA full insert					
sequencing; Research Association for Biotechnology; cDNA library					
construction; 5'- & 3'-end one pass sequencing and clone selection;					
Helix Research Institute (supported by Japan Key Technology Center					
etc.) and Department of Virology, Institute of Medical Science,					
University of Tokyo.					
FEATURES					
Location/Qualifiers					
1..1961					
/organism="Homo sapiens"					
/mol_type="mRNA"					
/db_xref="taxon:9606"					
/clone="NT2RP4000243"					
/cell_line="NT2"					
/cell_type="teratocarcinoma"					
/clone_lib="NT2RP4"					
/note="cloning vector: pME18SFL3-mRNA from NT2 neuronal					
precursor cells after 2-weeks retinoic acid (RA)					
induction."					
BASE COUNT	417 a	579 c	481 g	484 t	
ORIGIN					
Query Match					
Best Local Similarity					
Matches 1129;					
Conservative					
2;					
Mismatches					
3;					
Indels					
0;					
Gaps	0;				
QY	1	GGGCGCGCCCAATACAAAGCTTACAGCTTCCGAGCTTCCACGGGACGAGCTGATGCGG	60		
DB	109	GGGCGCGCCCAATACAAAGCTTACAGCTTCCGAGCTTCCACGGGACGAGCTGATGCGG	168		
QY	61	CTCGAGTCCGCTTACCGGACCGCTGGACAAGTACAGCGGCGAGCACTGGGCCGAGAGC	120		
DB	169	CTCGAGTCCGCTTACCGGACCGCTGGACAAGTACAGCGGCGAGCACTGGGCCGAGAGC	228		
QY	121	STKGGTACTCGGAGATCAGCTTCCGAGCTTCCACGGGACGAGCTTCCACGGGACGAGC	180		
DB	229	GTGGGCTACTCGGAGATCAGCTTCCGAGCTTCCACGGGACGAGCTTCCACGGGACGAGC	288		
QY	181	TGCCACCGCAACTGACGAGCGCGCGAGCGGCGAGCGCGCGCGCGCGCGCGCGCGCG	240		

```
Db 289 TGCACGCAACTGACAGCGCGCGCGAGCCGCGCGCTCGCAGCTAT 348
Qy 241 CCCGAGCTGCGCTCTTCGCGGGGCTGCTGCGCGCGCGCGCTCAAGCGCTGAG 300
Db 349 CCCGAGCTGCGCTCTTCGCGGGCTGCTGCGCGCGCGCGCTCAAGCGCTGAG 408
Qy 301 CAGGCGCTGCGAGCTTCGCGCACTCCAGCCAGCGCGAGGCTGCGCGGACTTCCAG 360
Db 409 CAGGCGCTGCGAGCTTCGCGCACTCCAGCCAGCGCGAGGCTGCGCGGACTTCCAG 468
Qy 361 CGCGCGAGCGCTCAAGTTCCTGAGTTCGCTTCAAGGCAATAATCTCCCAA 420
Db 469 CGCGCGAGCGCTCAAGTTCCTGAGTTCGCTTCAAGGCAATAATCTCCCAA 528
Qy 421 GCCATCGCGCTGCTCAAGTTCCTGAGTTCGCTTCAAGGCAATAATCTCCCAA 480
Db 529 GCCATCGCGCTGCTCAAGTTCCTGAGTTCGCTTCAAGGCAATAATCTCCCAA 588
Qy 481 AACATGCGATATTAAGAGCTTCGCTGAGTTCGCTTCAAGGCAATAATCTCCCAA 540
Db 589 AACATGCGATATTAAGAGCTTCGCTGAGTTCGCTTCAAGGCAATAATCTCCCAA 648
Qy 541 AAGTCATATGAAGCTTCGCTGAGTTCGCTTCAAGGCAATAATCTCCCAA 600
Db 649 AAGTCATATGAAGCTTCGCTGAGTTCGCTTCAAGGCAATAATCTCCCAA 708
Qy 601 ACATCCATCAGACATGAGCTTCGCTGAGTTCGCTTCAAGGCAATAATCTCCCAA 660
Db 709 ACATCCATCAGACATGAGCTTCGCTGAGTTCGCTTCAAGGCAATAATCTCCCAA 768
Qy 661 CTCGAGCTCGAGGCTTCGAGGAGATCAAGAGCTTCGAGGATTTTACGAGTGT 720
Db 769 CTCGAGCTCGAGGCTTCGAGGAGATCAAGAGCTTCGAGGATTTTACGAGTGT 828
Qy 721 GCAGATCATATGAGAGTTCGAGGAGATCAAGAGCTTCGAGGATTTTACGAGTGT 780
Db 829 GCAGATCATATGAGAGTTCGAGGAGATCAAGAGCTTCGAGGATTTTACGAGTGT 888
Qy 781 GTTATGAGAGCTATCCGTTTGAAGATTTGCTGCTACCATGATATTAATCTGAGTTT 840
Db 889 GTTATGAGAGCTATCCGTTTGAAGATTTGCTGCTACCATGATATTAATCTGAGTTT 948
Qy 841 GCCTATATAGTTCGAGGAGCTTCGAGGAGATCAAGAGCTTCGAGGATTTTACGAGTGT 900
Db 949 GCCTATATAGTTCGAGGAGCTTCGAGGAGATCAAGAGCTTCGAGGATTTTACGAGTGT 1008
Qy 901 TTTGATCAGATGCAAGGCTCATGAGAGCTTCGAGGAGATCAAGAGCTTCGAGGAGCT 960
Db 1009 TTTGATCAGATGCAAGGCTCATGAGAGCTTCGAGGAGATCAAGAGCTTCGAGGAGCT 1068
Qy 961 ACTTGGGCTCTCGAGTTCGAGGAGCTTCGAGGAGATCAAGAGCTTCGAGGAGCTTCGAGT 1020
Db 1069 ACTTGGGCTCTCGAGTTCGAGGAGCTTCGAGGAGATCAAGAGCTTCGAGGAGCTTCGAGT 1128
Qy 1021 GTGACCACTCCAGAGAGCTGTATGAGTTCGAGGAGATCAAGAGCTTCGAGGAGCTTCGAGT 1080
Db 1129 GTGACCACTCCAGAGAGCTGTATGAGTTCGAGGAGATCAAGAGCTTCGAGGAGCTTCGAGT 1188
Qy 1081 GAGGAGAGTTCGAGGAGATCAAGAGCTTCGAGGAGATCAAGAGCTTCGAGGAGCTTCGAGT 1134
Db 1189 GAGGAGAGTTCGAGGAGATCAAGAGCTTCGAGGAGATCAAGAGCTTCGAGGAGCTTCGAGT 1242
```

RESULT 6

```
AK075401
LOCUS AK075401 2010 bp mRNA linear PRI 03-SEP-2002
DEFINITION Homo sapiens cDNA PSEC0091 f15, clone PLACE1001683, highly similar
to Homo sapiens mRNA for cartilage-associated protein.
ACCESSION AK075401
VERSION AK075401.1 GI:22761465
KEYWORDS oligo capping, f15 (full insert sequence).
SOURCE Homo sapiens (human)
```

```
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Ota, T., Nishikawa, T., Suzuki, Y., Kawai-Hio, Y., Hayashi, K.,
JOURNAL Ishii, S., Saito, K., Yamamoto, J., Wakamatsu, A., Nagai, T.,
REFERENCE Nakamura, Y., Sato, H., Nagahara, K., Sugano, S. and Isogai, T.
AUTHORS HRI human cDNA sequencing project
TITLE HRI human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2010)
AUTHORS Isogai, T. and Yamamoto, J.
TITLE Direct Submission
JOURNAL Submitted (20-MAR-2002) Takao Isogai, Helix Research Institute,
COMMENT Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
HRI human cDNA sequencing project; cDNA 5' - 3' - end one pass
sequencing, clone selection and full insert sequencing; Helix
Research Institute (supported by Japan Key Technology Center etc.);
cDNA library construction; Institute of Medical Science, University
of Tokyo, Laboratory of Genome Structure, Human Genome Center.
FEATURES
Location/Qualifiers
1..2010
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE1001683"
/tissue_type="placenta"
/clone_lib="PLACE1"
/note="Cloning vector: pME18SFL3"
misc_difference 728*729
/note="compared to AK001634 and AJ006470"
/replace="c"
BASE COUNT 416 a 606 c 485 g 503 t
ORIGIN
Query Match 98.6%; Score 1118; DB 9; Length 2010;
Best Local Similarity 99.6%; Pred. No. 8.7e-178;
Matches 1129; Conservative 2; Mismatches 2; Indels 1; Gaps 1;
Qy 1 GGGCGCGCCCAATACGAAGCTTACAGCTTCGCGAGCTTCCCGAGGAGCGAGCTGATGCGG 60
Db 165 GGGCGCGCCCAATACGAAGCTTACAGCTTCGCGAGCTTCCCGAGGAGCGAGCTGATGCGG 224
Qy 61 CTCGAGTGGCGCTACCGGACGCGCTGGACAAGTACAGCGCGGAGCACTGGCGCGAGAGC 120
Db 225 CTCGAGTGGCGCTACCGGACGCGCTGGACAAGTACAGCGCGGAGCACTGGCGCGAGAGC 284
Qy 121 STKGCTACCTCGAGATCAGCTTCGCGGTGCAACCGCTTGTGCGCGAGCAGCGAGCTTC 180
Db 285 GTGGGCTACCTCGAGATCAGCTTCGCGGTGCAACCGCTTGTGCGCGAGCAGCGAGCTTC 344
Qy 181 TGCCACCGCAACTGCGAGCGCGCGCGGAGCGCGGAGCGCGGAGCGCGCTCGCCAGCTAT 240
Db 345 TGCCACCGCAACTGCGAGCGCGCGCGGAGCGCGGAGCGCGGAGCGCGCTCGCCAGCTAT 404
Qy 241 CCGGAGCTCGGCTCTTTCGGGGGCTGCTGCGCGCGCGGAGCGCGGAGCGCGCTCAAG 300
Db 405 CCGGAGCTCGGCTCTTTCGGGGGCTGCTGCGCGCGCGGAGCGCGGAGCGCGCTCAAG 464
Qy 301 CAGGCGCTCGGAGCTTCGCGGAGTCCAGCCAGTCCAGCCAGCGAGGCTGCGGAGCTTC 360
Db 465 CAGGCGCTCGGAGCTTCGCGGAGTCCAGCCAGTCCAGCCAGCGAGGCTGCGGAGCTTC 524
Qy 361 CGCGCGAGCGCTTACGAAGTTCCTGAGTTCGCTTACTTCAAGGCAATAATCTCCCAA 420
Db 525 CGCGCGAGCGCTTACGAAGTTCCTGAGTTCGCTTACTTCAAGGCAATAATCTCCCAA 584
Qy 421 GCATCGCGCTGCTCAGACCTTCTTCTGAGAGCATCTGATGAGCAATAATGTAAGAGG 480
Db 585 GCATCGCGCTGCTCAGACCTTCTTCTGAGAGCATCTGATGAGCAATAATGTAAGAGG 644
Qy 481 ACATGCGATATTAAGAGCTTCGCTGAGTTCGCTTCAAGGCAATAATCTCCCAA 540
```

Db 645 AACATGGCATATTATTAAGACCTGCTGGTGGCGGAGGACTACATTAAAGACCTGGAACC 704
QY 541 AAGTATATGAAGCCTGTTATCCAGCAGTGGCGGCATACAAAGCTGAGAACTGGAGA 600
Db 705 AAGTCATATGAAGCCTGTTAT-CGAGCAGTGGCGGCATACAAAGCTGAGAACTGGAGA 763
QY 601 ACATCCATCACACATGAGCTGGCCCTCCGACTTCTTCAAGCCCTTTTACAGTGT 660
Db 764 ACATCCATCACACATGAGCTGGCCCTCCGACTTCTTCAAGCCCTTTTACAGTGT 823
QY 661 CTGGAGCCTGGAGGCTTCCAGGAGATCAAGGACTTCAAGATTTTACCTTCCATA 720
Db 824 CTGGAGCCTGGAGGCTTCCAGGAGATCAAGGACTTCAAGATTTTACCTTCCATA 883
QY 721 GCAGATCATTATGTAAGTCTGGAATGCAAAATACAGTGTGAAGAGAACTCACCCCA 780
Db 884 GCAGATCATTATGTAAGTCTGGAATGCAAAATACAGTGTGAAGAGAACTCACCCCA 943
QY 781 GTTATAGGAGGCTATCCGTTGAGAAATTTGGCTACCATGATCATTTACCTGAGTGT 840
Db 944 GTTATAGGAGGCTATCCGTTGAGAAATTTGGCTACCATGATCATTTACCTGAGTGT 1003
QY 841 GCTATATATAGTGAAGCCTGAGATGAGAGCCCTGTCAGTCAAGTATCTGCTC 900
Db 1004 GCTATATATAGTGAAGCCTGAGATGAGAGCCCTGTCAGTCAAGTATCTGCTC 1063
QY 901 TTTGATCAGAAATGACAGGCTCATGAGCAGAACTGCTGTTATACAGTACCACAGGAC 960
Db 1064 TTTGATCAGAAATGACAGGCTCATGAGCAGAACTGCTGTTATACAGTACCACAGGAC 1123
QY 961 ACTTGGGCTCTCGGATGAGCACTTCCAGCCAGACCTGAGCAGTTCAGTTCITTAAT 1020
Db 1124 ACTTGGGCTCTCGGATGAGCACTTCCAGCCAGACCTGAGCAGTTCAGTTCITTAAT 1183
QY 1021 GTGACCACTTCCAGAGGAGCTGATGACTTTGCTAAGGAAATATAATGATGATGAT 1080
Db 1184 GTGACCACTTCCAGAGGAGCTGATGACTTTGCTAAGGAAATATAATGATGATGAT 1243
QY 1081 GAGGAGAAATGTTGGAAATGTTGATGAGCTTCTTGAATGAGAGACCAAC 1134
Db 1244 GAGGAGAAATGTTGGAAATGTTGATGAGCTTCTTGAATGAGAGACCAAC 1297

RESULT 7

AK091772 1922 bp mRNA linear PRI 15-JUL-2002
LOCUS Homo sapiens cDNA FLJ34453 fis, clone HLUNG2002429, highly similar
DEFINITION to Homo sapiens mRNA for cartilage-associated protein.

AK091772
ACCESSION AK091772.1 GI:21750221

KEYWORDS oligo capping; fis (full insert sequence).

SOURCE Homo sapiens

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 Tashiro, H., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S.,
Fukuzumi, Y., Fujimori, Y., Komiyama, M., Sugiyama, T., Irie, R.,
Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y.,
Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H.,
Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,
Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A.,
Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,
Nagahari, K., Masuho, Y., Nagai, K. and Isogai, T.
NEDO human cDNA sequencing project
Unpublished

2 (bases 1 to 1922)

Isogai, T. and Yamamoto, J.

Direct Submission

AUTHORS

Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)

JOURNAL

NEDO human cDNA sequencing project supported by Ministry of

Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.

FEATURES

source

Location/Qualifiers
1..1922
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HLUNG2002429"
/tissue_type="lung"
/clone_lib="HLUNG2"
/notes="cloning vector: pME18SFL3"
40..1206
/note="unnamed protein product"
/codon_start=1
/protein_id="BAC03743.1"
/db_xref="GI:21750222"
/translation="MEFRRGAALLALLCVACALRAGRAYERYSPRPRDLMLPL
ESAVRHLDKYSGEHWAESVRDEAFCHRNCSAOPPEAAGLASPELFGDLRLR
AHCLRKQGLPAFROSQPSREVLADFQREPKFLQFAYFANNLPKATAAHTFL
KHPDEMKNMAYKSLPCAEDYIKDLETKYESLFRVAVRYNENMETITDMEL
ALDPFFAFYECLEAACGSEIKDFDKFYLSDHVEVELECKIENETPVIGYVP
VERFVATWYHLYQFAYKLNLDKNAAPCAVSYLLFDONDKVMOONLVYVYHRTWEL
SDEHFQRPRAVQFFNVTTLQKELYINDDDEGEVEVEYVDDLLEETS"

CDS

BASE COUNT 413 a 566 c 467 g 476 t
ORIGIN
Query Match 91.9%; Score 1042.4; DB 9; Length 1922;
Best Local Similarity 96.3%; Pred. No. 4.2e-165;
Matches 1092; Conservative 1; Mismatches 2; Indels 39; Gaps 1;
QY 1 GGGCGCGCCCAATACGAACGCTCAGCTTCGCGAGCTTCCACCGGACGAGCTGATCGG 60
Db 109 GGGCGCGCCCAATACGAACGCTCAGCTTCGCGAGCTTCCACCGGACGAGCTGATCGG 168
QY 61 CTCGAGTCGGCTACCGGACGCGCTGGACAAGTACAGCGCGGAGGAGCTGGCGCGAGAGC 120
Db 169 CTCGAGTCGGCTACCGGACGCGCTGGACAAGTACAGCGCGGAGGAGCTGGCGCGAGAGC 228
QY 121 STKGGCTACCTGGAGATCAGCGCTCGGGCTGCGCCGCTTGCTTGGCGGACGAGGCGCTTC 180
Db 229 G-----TGCGGACGAGGAGGCGCTTC 249
QY 181 TGCCACCGCAACTGTCAGCGCGCGCGGACGCGCGCGCGCGCGCTTCGCCAGCTAT 240
Db 250 TGCCACCGCAACTGTCAGCGCGCGCGGACGCGCGCGCGCGCGCTTCGCCAGCTAT 309
QY 241 CCGGAGCTCGGCTCTTCGGGGGCTGCTGGCGCGCGGAGCTGCTCAAGCGTGCAG 300
Db 310 CCGGAGCTCGGCTCTTCGGGGGCTGCTGGCGCGCGGAGCTGCTCAAGCGTGCAG 369
QY 301 CAGGCGCTCGGAGCTTCGGCGAGTCCAGCGCGGAGGCTGCTGGCGAGCTTCAG 360
Db 370 CAGGCGCTCGGAGCTTCGGCGAGTCCAGCGCGGAGGCTGCTGGCGAGCTTCAG 429
QY 361 CGCGCGGAGCTTCAAGTTCCTGAGTTCCTTCAAGGCAATAATCTCCCAAAA 420
Db 430 CGCGCGGAGCTTCAAGTTCCTGAGTTCCTTCAAGGCAATAATCTCCCAAAA 489
QY 421 GCCATCGCGCTGCTCACACCTTTTCTACTGAAGCATCTCTGATGAGAAATGATGAAGG 480
Db 490 GCCATCGCGCTGCTCACACCTTTTCTACTGAAGCATCTCTGATGAGAAATGATGAAGG 549
QY 481 AACATGGCATATTATTAAGAGCCTGCTGGTCCGAGGAGTACATTAAAGACTGGAACC 540
Db 550 AACATGGCATATTATTAAGAGCCTGCTGGTCCGAGGAGTACATTAAAGACTGGAACC 609
QY 541 AAGTCATATGAAGGCTGTTCATCCGAGCAGTGGCGGCATACAAAGCTGAGAACTGGAGA 600

Db 610 AAGTCATATGAAGCCTGTTCTACCGAGCAGTGGCGGCATACACGGTGAGAACTGGAGA 669
 Qy 601 ACATCCATCACAGACATGAGCTGGCCCTTCCCGACTTCTTCAAAGCCTTTTACGAGTGT 660
 Db 670 ACATCCATCACAGACATGAGCTGGCCCTTCCCGACTTCTTCAAAGCCTTTTACGAGTGT 729
 Qy 661 CTCGCGAGCTGGAGGGTTCAGGAGATCAGGACTTCAAGGATTTCTACCTTTCCATA 720
 Db 730 CTCGCGAGCTGGAGGGTTCAGGAGATCAGGACTTCAAGGATTTCTACCTTTCCATA 789
 Qy 721 CGAGATCATATGTAGAAGTCTGGAATGCAAAATACAGTGTGAAGAACTTCAACCCCA 780
 Db 790 GCAGATCATATGTAGAAGTCTGGAATGCAAAATACAGTGTGAAGAACTTCAACCCCA 849
 Qy 781 GTTATAGGAGGCTATCCGGTTCAGAAATTTGGGCTACCATGTATCTATCTTCAGTTT 840
 Db 850 GTTATAGGAGGCTATCCGGTTCAGAAATTTGGGCTACCATGTATCTATCTTCAGTTT 909
 Qy 841 GCCTATTATAGTTGAACCTGGAAGAAATGCGAGCCCTCTGCGAGTCAGCTATCTGCTC 900
 Db 910 GCCTATTATAGTTGAACCTGGAAGAAATGCGAGCCCTCTGCGAGTCAGCTATCTGCTC 969
 Qy 901 TTGTATCAGATGACAAAGTCTATGAGCAGAACCTGCTGTTATACCGATGCCAGGGAC 960
 Db 970 TTGTATCAGATGACAAAGTCTATGAGCAGAACCTGCTGTTATACCGATGCCAGGGAC 1029
 Qy 961 ACTTGGGGCTCTCGATGAGCACTTCAGCCAGACCTGAAGCAGTTCAGTTCTTTAAT 1020
 Db 1030 ACTTGGGGCTCTCGATGAGCACTTCAGCCAGACCTGAAGCAGTTCAGTTCTTTAAT 1089
 Qy 1021 GTGACCACACTCCAGAAAGAGCTGTATGACTTTGCTAAGGAAATATATATGATGATGAT 1080
 Db 1090 GTGACCACACTCCAGAAAGAGCTGTATGACTTTGCTAAGGAAATATATATGATGATGAT 1149
 Qy 1081 GAGGGAAGATGTGGAATATGTGATGACCTTTGGAACCTGGAGGAGACCCAGC 1134
 Db 1150 GAGGGAAGATGTGGAATATGTGATGACCTTTGGAACCTGGAGGAGACCCAGC 1203

RESULT 8
 AK092990 1903 bp mRNA linear PRI 15-JUL-2002
 LOCUS Homo sapiens cDNA FLJ35671 fis, clone SPLEN2018180, highly similar
 DEFINITION to Homo sapiens mRNA for cartilage-associated protein.
 AK092990
 ACCESSION AK092990.1 GI:21751708
 VERSION
 KEYWORDS oligo capping; fis (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 1 Tashiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Ota,T., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
 NEDO human cDNA sequencing project
 Unpublished
 2 (bases 1 to 1903)
 Isogai,T. and Yamamoto,J.
 Direct Submission
 TITLE Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamata, Kisarazu, Chiba 292-0812, Japan
 (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
 NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,

HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.
 FEATURES
 Location/Qualifiers
 1. 1903
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="SPLEN2018180"
 /tissue_type="spleen"
 /clone_lib="SPLEN2"
 /note="cloning vector: pME18SFL3"
 BASE COUNT 421 a 553 c 460 g 469 t
 ORIGIN
 Query Match 89.9%; Score 1019.6; DB 9; Length 1903;
 Best Local Similarity 99.7%; Pred. No. 2.8e-161;
 Matches 1018; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 Qy 114 CGAGAGCTGCGCTACCTGGAGATCAGCTTGGCGCTGCACCGTTGCTGCGGACAGCA 173
 Db 147 CGAGAGCTGCGCTACCTGGAGATCAGCTTGGCGCTGCACCGTTGCTGCGGACAGCA 206
 Qy 174 GGCCTTCTCCACCGCAACTGCAGCGCGCGCGAGCCCGCGCGCGCGCGCGCTCGC 233
 Db 207 GGCCTTCTCCACCGCAACTGCAGCGCGCGCGAGCCCGCGCGCGCGCGCGCTCGC 266
 Qy 234 CAGCTATCCGAGCTGGCGCTCTTTCGGGGGCGCTGCTGCGCGCGCGCACTTCAAGCG 293
 Db 267 CAGCTATCCGAGCTGGCGCTCTTTCGGGGGCGCTGCTGCGCGCGCGCACTTCAAGCG 326
 Qy 294 CTGCAAGCAGCGCTGCGCGCTTCCGCGAGTCCCGAGCCCGCGCGCGCGCGCTGCGCGCA 353
 Db 327 CTGCAAGCAGCGCTGCGCGCTTCCCGAGTCCCGAGCCCGCGCGCGCGCGCTGCGCGCA 386
 Qy 354 CTTCCAGCGCGCGAGCCCTACAGTTTCCTGCGAGTTGCTTACTTCAAGCAAAATATCT 413
 Db 387 CTTCCAGCGCGCGAGCCCTACAGTTTCCTGCGAGTTGCTTACTTCAAGCAAAATATCT 446
 Qy 414 CCCCAAGCCATCGCGCTGCTCACACCTTCTTACTGAGCATCTCTGATGAGCAATGAT 473
 Db 447 CCCCAAGCCATCGCGCTGCTCACACCTTCTTACTGAGCATCTCTGATGAGCAATGAT 506
 Qy 474 GAAGAGCAATGCGCATATTATAAGACCTGCTGCTGCGCGAGCTTACATTAAAGACCT 533
 Db 507 GAAGAGCAATGCGCATATTATAAGACCTGCTGCTGCGCGAGCTTACATTAAAGACCT 566
 Qy 534 GGAACCAAGTCATATGAAGCCTGTTTCATCCGAGCAGTCCGCGCATACACGTCAGAA 593
 Db 567 GGAACCAAGTCATATGAAGCCTGTTTCATCCGAGCAGTCCGCGCATACACGTCAGAA 626
 Qy 594 CTGGAGAACATCCATCAGACATGGAGCTGGCGCTTCCGAGCTTCTTCAAGCCTTTTA 653
 Db 627 CTGGAGAACATCCATCAGACATGGAGCTGGCGCTTCCGAGCTTCTTCAAGCCTTTTA 686
 Qy 654 CGAGTGTCTCCAGCGCTGCGAGGGTTCAGGAGATCAGGACTTCAAGGATTTTACCT 713
 Db 687 CGAGTGTCTCCAGCGCTGCGAGGGTTCAGGAGATCAGGACTTCAAGGATTTTACCT 746
 Qy 714 TTCCATAGCAGATCATTATGTAGAGTTCTGGAATGCAAAATACAGTGTGAAGAACCT 773
 Db 747 TTCCATAGCAGATCATTATGTAGAGTTCTGGAATGCAAAATACAGTGTGAAGAACCT 806
 Qy 774 CACCCCGATTATAGAGGCTATCCGGTTGAGAAATTTGGCTACCATCTATCATTTACT 833
 Db 807 CACCCCGATTATAGAGGCTATCCGGTTGAGAAATTTGGCTACCATCTATCATTTACT 866
 Qy 834 GCAGTTTCCCTATTATTAAGTTGAACACCTGGAAGATGAGACCCCTGTGCGAGTACGTA 893
 Db 867 GCAGTTTCCCTATTATTAAGTTGAACACCTGGAAGATGAGACCCCTGTGCGAGTACGTA 926
 Qy 894 TCTGCTCTTTCATCAGAAATCAGAGGTCATGCGAGCAGAACCTGGTGTATTACCAGTACCA 953

FEATURES	SOURCE
1..1663	/organism="Mus musculus"
1..1663	/mol_type="mRNA"
1..1663	/db_xref="taxon:10090"
1..1663	/clone="MGC:54548 IMAGE:6306169"
1..1663	/tissue_type="Olfactory epithelium, neonatal mouse, C57Bl/6"
1..1663	/clone_lib="NIH MGC_129"
1..1663	/lab_host="DH10B"
1..1663	/note="Vector: PCMV-SPORT6"
1..1663	/genes="Crtap"
1..1663	/notes="synonyms: CASP, 57305292N3rik"
1..1663	/db_xref="LocustID:56693"
1..1663	/db_xref="MGI:1891221"
1..1663	51..1253
1..1663	/codon_start=1
1..1663	/product="cartilage associated protein"
1..1663	/protein_id="AAH43890.1"
1..1663	/db_xref="GI:29437151"
1..1663	/db_xref="LocusID:56693"
1..1663	/translation="MGPSRPTAALLVLCVCAPTGRGOYERYSRSPFDRLMPLE SAYRLADDOYSGEHWAEVGVLESLRHLRLDRSEAFCHRNCSAATPAFAGPASH AEURLGVSRQAQCLGAFQROSPSLADFOREPXYFLFPAFYKANDL PKAIAAHTYLLKHPDDMKWMEVYKSLPGABDHKLDTKSYESLFPVAVKAYNG ENKRTSISMEWALPDFKAFCEVLAACEGSREIKDFKDFYLSIADHYVEVLECKRC FEYLTAVIGYGVPEKVFATVHYVLFQAYKLLDNKNAAPCAVSVLLFDOSDRWQOOL VYVQYHRDKNGLSDEHPRPEAVQFNVTTLQKELYDFAQERHLMDDDSGEVVEYVDD LLETSSA"
BASE COUNT	370 a 490 c 442 g 361 t
ORIGIN	
Query Match	79.1%; Score 897.2; DB 10; Length 1663;
Best Local Similarity	87.0%; Pred. No. 9, 4e-141;
Matches 983; Conservative	2; Mismatches 145; Indels 0; Gaps 0;
QY	1 GGGCGCGCCCAATACGAAGCTACAGCTTCGCGAGCTTCCACGGGACGAGCTGATGCCG 60
DB	117 GGGCGCGGGCAGTACAGAGCGCTACAGCTTCGCGAGCTTCGCGGGACGAGCTGATGCCG 176
QY	61 CTCAGGTGCGGCTTACCGGCACGCGCTGGACAAGTACAGCGCGGACGACTGGCGCGAGAGC 120
DB	177 CTCAGGTGCGGCTTACCGGCACGCGCTTCGCCAGTACAGCGCGGACGACTGGCGCGAGAGC 236
QY	121 STKGCTACCTGAGATACGCTTGGCGGTGACCGCTTGTCGCCGACACGACGAGGCTTC 180
DB	237 GTGGGCTACTCTGAGGTGAGCTTCGGGTGACCGCTGCTGGCGCAGCAGGAGGCTTC 296
QY	181 TGCCACCGCAACTGACGAGCGCGCGCAGCGCAGCGCGCGCGGCTTCGCCAGCTAT 240
DB	297 TGCCACCGCAACTGACGAGCGCGCCACGCGCAGCACC CGCGCGCGCGCGCGCCAC 356
QY	241 CCGAGCTGGCGCTCTTGGGGGCTGTGTGGCGCGCGGACCTGCTCTAAGCGCTCGAG 300
DB	357 GCCGAATACGCTCTTCGGGACGCTGTGTGGCGCGCGCGCAGTGCCTCAAGCGCTCGAAG 416
QY	301 CAGGCGCTGCGACGCTTCGCCCGCAGTCCCGCAGCGCGCGCGCGCGCTTCGCCAG 360


```

Db      417 CAGGGCTGCGCGCTTCGCCAGTCCGACGCCGCTCAGTCTGGCGGACTTTTCAG 476
Qy      361 CGCGGAGCCCTACAGTTCCTGCGATTCCTTCAAGGCAATTAATCTCCCAAA 420
Db      477 CAGCGGAGCCCTACAGTTCCTGCGATTCCTTCAAGGCAATTAATCTCCCAAA 536
Qy      421 GCATCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Db      537 GCATCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 596
Qy      481 AACATGCAATTAATTAAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Db      597 AACATGCAATTAATTAAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 656
Qy      541 AAGTCATATGAAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Db      657 AAGTCGATGAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 716
Qy      601 ACATCCATCACAGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db      717 ACATCCATCACAGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 776
Qy      661 CTGCGAGCCCTGCGAGGTTCCAGGAGATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGAT 720
Db      777 CTGCGAGCCCTGCGAGGTTCCAGGAGATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGAT 836
Qy      721 GCAGATCAATTAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db      837 GCAGATCAATTAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 896
Qy      781 GTTATAGGAGCTATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db      897 GTTATAGGAGCTATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 956
Qy      841 GCTTATTAAGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db      957 GCTTATTAAGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1016
Qy      901 TTTGATCAGATGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Db      1017 TTTGATCAGATGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1076
Qy      961 ACTTGGGGCTCTCGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Db      1077 AAGTGGGGCTCTCGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1136
Qy      1021 GTGACCACTCTCGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Db      1137 GTGACCACTCTCGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1196
Qy      1081 GAGGGAGAGTGTGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1130
Db      1197 GAGGGAGAGTGTGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1246

RESULT 10
MMAJ6469 1695 bp mRNA linear ROD 02-AUG-1999
LOCUS Mus musculus mRNA for cartilage-associated protein (CASP).
DEFINITION
ACCESSION AJ006469
VERSION AJ006469.1 GI:3687319
KEYWORDS cartilage-associated protein; CASP.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
1 Morello, R., Tonachini, L., Monticone, M., Viggiano, L., Rocchi, M.,
Cancedda, R. and Castagnola, P.
cDNA cloning, characterization and chromosome mapping of Crtp
encoding the mouse cartilage associated protein
JOURNAL Matrix Biol. 18 (3), 319-324 (1999)
MEDLINE 99357019

```

```

PUBMED 10429950
REFERENCE 2 (bases 1 to 1695)
AUTHORS Castagnola, P.
TITLES Direct Submission
JOURNAL Submitted (12-JUN-1998) Castagnola P., Biotechnologie in Oncologia,
Ist. Naz.le per la Ricerca sul Cancro, L.go R. Benzi 10 Genoa,
16132, ITALY
FEATURES
Location/Qualifiers
source 1..1695
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="embryo"
/dev_stage="14-18 d.p.c."
gene 1..1695
/genes="casp"
CDS 72..1274
/genes="casp"
/codon_start=1
/product="cartilage-associated protein (CASP)"
/protein_id="CAA07053.1"
/db_xref="GI:3687320"
/db_xref="SPTRMSL:088698"
/translation="MGPRSPAAALLVLLCVGCAPTGRCQYERYSFRNPPDELMPLD
AEJLFGSVLRRAQCLKRCQGLPAPFQSPQSRSLADFOQREPYKFLQFAYFKANDL
PKATAAHTYLLKHPDQDEMMKEMMYKSLPGAEDHIKDLBTKSYESLFAVAYKNG
ENWRTSISDMELALPDLFKAFYECIAACEGSRDKDFKDFYLSIAHDFVLECKIRC
EETLTPVIGYVPEKFTVWYHLOFAYVYKLNLDKNAAPCAVSYLLPDQSRVMOQNL
VYCYHRDKWGLSDEHFQPRPEAVQFNVTTLQKELYDFAQEHLMDDDEGEVVEYDD
LLETFESA"
sig_peptide 72..146
/mat_peptide 147..1271
BASE COUNT 362 a 503 c 449 g 371 t
ORIGIN
Query Match 79.08; Score 895.6; DB 10; Length 1695;
Best Local Similarity 86.99; Pred. No. 1.7e-140; Indels 0; Gaps 0;
Matches 982; Conservative 2; Mismatches 146;
Qy 1 GGGCGCGCCCAATAAGAACCTACAGCTTCCGAGCTTCCACGGAGCAGAGCTGATGCCG 60
Db 138 GGGCGCGGGCAGTACGAGCGCTACAGCTTCCGCACTTCCCGCGGAGCAGCTGATGCCG 197
Qy 61 CTCGAGTCGCGCTACCGGACGCGCTGGACAAGTACAGCGGCGGAGCAGCTGGCGCGAGAGC 120
Db 198 CTCGAGTCGCGCTACCGGACGCGCTTCGACAGTACAGCGGCGGAGCAGCTGGCGCGAGAGC 257
Qy 121 STKGCTACCTGGAGATCAGCTTGGCGCTGCAACCGCTTGTCTGGCGGAGCAGCGAGCCTTC 180
Db 258 GTGGGCTACCTGGAGGAGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 317
Qy 181 TGCCACCGCAACTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCGCGAGCTAT 240
Db 318 TGCCACCGCAACTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 377
Qy 241 CCCGAGCTGCGCTTCTTGGGGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
Db 378 GCCGAATACGCTCTTTCGCGAGCGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 437
Qy 301 CAGGGCTGCGAGCGCTTCCGCCAGTCCCGAGCAGCGCGCGCGCGCGCGCGCGCGCGCGAG 360
Db 438 CAGGGCTGCGCGCTTCCGCCAGTCCCGAGCAGCGCGCTCAGTCTGCGCGAGCTTTTCAG 497
Qy 361 CGCGCGGAGCCCTACAGTTCCTGCGAGTTCGCTTCAAGGCAATTAATCTCCCAAA 420
Db 498 CAGCGGAGCCCTACAGTTCCTGCGAGTTCGCTTCAAGGCAATTAATCTCCCAAA 557
Qy 421 GCCATCGCGCTGCTCAGACCTTTCTTACTGAAGCATCTCTGATGACGAATGATGAAGAGG 480

```


Db 558 GCATCGCTCGGCTCACACCTATCTCTGAGCATCCAGATGACGAGATGAAGAGA 617
 QY 481 AACATGGCATATATAAGAGCCTGCTGTCGCCGAGGACTACATTAAGACCTGGAAC 540
 Db 618 AACATGGAGTATATAAGAGCTTGCTGAGCCGAGGACCAACATTAAGACTTGGAA 677
 QY 541 AAGTCATATGAAGCTGTTCATCCGAGCAGTGGGCGATACACGCTGGAACCTGGAG 600
 Db 678 AAGTCGACGAGCCTGTTCTCCGTCGGTGCGGCTCAACGCGGAGAACTGGAGA 737
 QY 601 ACATCCATCACAGACATGAGCTGCCCTTCCGACTTCTTCAAGCTTTTACGAGTGT 660
 Db 738 AGCTCCATTTCCGACATGAGCTCGGCTTCCGACTTCTTCAAGGCTTCTACGAGTGC 797
 QY 661 CTCGAGCTCGAGGCTTCCAGGAGATCAGGACTTCAAGGATTTCTACCTTTCCATA 720
 Db 798 CTGGCTGCTCGAGGCTCGCGGAGATCAAGACTTCAAGGACTTCTACCTTGTCCATA 857
 QY 721 CGAGATCATATGTAGAAGTCTGGAATGCAAAATACAGTGTGAAGAGAACTCACCCCA 780
 Db 858 GCAGATCACTATGTGAAGTCTGGAGTGAAGATTCGTTGTGAGGAGACCTCACCCCA 917
 QY 781 GTTATAGAGGCTATCCGCTTGAGAAATTTGCTGCTACCATGTATCATTTTCGAGTT 840
 Db 918 GTCATAGAGGCTATCCGCTGAGAAATTTGCGGACCATGTACCACTATTACAGTTT 977
 QY 841 GCCTATTATTAAGTTGAACGACTGAAGAAATGACGCCCTGTGCACTCAGCTATCTGCTC 900
 Db 978 GCGTATTACAGTTGAATGATCTGAAGAAATGACGCCCTGTGCGCTCAGCTACCTGCTC 1037
 QY 901 TTTGATCAGAAATGACAGGCTATGACAGAACTGCTGATTTACCACTACCACTACAGGAC 960
 Db 1038 TTTGACCAAGTACAGGCTATGACAGAACTGCTGATTTACCACTACCACTACAGGAC 1097
 QY 961 ACTTGGGCTCTCGGATGAGCACTTCCAGCCGAGACCTGAAGCACTTCACTTCTTTAAT 1020
 Db 1098 AAGTGGGCTCTCGGATGAGCACTTCCAGCCGAGACCCGAGCACTTCACTTCTTTAAT 1157
 QY 1021 GTGACCACTCCAGAGGCTGATGACTTCTTAAGGAATATTAAGTATGAT 1080
 Db 1158 GTGACAGCTCCAGAGGAACTGACACTTCTCAGAACACCTTGAAGATGAGAT 1217
 QY 1081 GAGGAGAGTTGTGGAATATGTGATGACCTTGGAACTGGAGGAGAC 1130
 Db 1218 GAGGAGAGTTGTGAGTATGTGAGCACTTGTGAGACGGAAGATC 1267

RESULT 11

AK097650

LOCUS

DEFINITION

Homo sapiens cDNA FLJ40331 fis, clone TEST12031687, highly similar
 to Homo sapiens mRNA for cartilage-associated protein.

ACCESSION

AK097650.1

VERSION

GI:21757488

KEYWORDS

oligo capping; fis (full insert sequence).

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

REFERENCE

1

Tanai H., Watanabe S., Ishida S., Ono Y., Horita T., Watanabe M.,

Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,

Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,

Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,

Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A.,

Kawakami B., Suzuki Y., Sugano S., Nagahara K., Masuho Y., Nagai K.

and Isogai T.

NEDO human cDNA sequencing project

Unpublished

2 (bases 1 to 1810)

Isogai T. and Yamamoto J.

Direct Submission

JOURNAL

Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
 Karusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
 (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)

COMMENT

NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:
 Research Association for Biotechnology (RAB); cDNA library
 construction: Helix Research Institute (HRI) (supported by Japan
 Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
 HRI, and Biotechnology Center, National Institute of Technology and
 Evaluation; clone selection for full insert sequencing: HRI and
 RAB; annotation: HRI and RAB.

FEATURES

source

1. 1810

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="TEST12031687"

/tissue_type="testis"

/clone_lib="TEST12"

/note="cloning vector: pMB18SPL3"

375 a 550 c 449 g 436 t

BASE COUNT

ORIGIN

Query Match 75.0%; Score 850; DB 9; Length 1810;
 Best Local Similarity 88.2%; Pred. No. 7.7e-133;
 Matches 1000; Conservative 2; Mismatches 2; Indels 130; Gaps 2;
 QY 1 GGGCGCGCCCAATACGAACGCTACAGCTTCCGAGCTTCCACGGGACGAGCTGATGCCG 60
 Db 94 GGGCGCGCCCAATACGAACGCTACAGCTTCCGAGCTTCCACGGGACGAGCTGATGCCG 153
 QY 61 CTGAGTCTGGCTACCGGCGAGCGCTGGACAAAGTACAGCGGCGAGCACTGGGCGGAGAGC 120
 Db 154 CTCGAGTCTGGCTACCGGCGAGCGCTGGACAAAGTACAGCGGCGAGCACTGGGCGGAGAGC 213
 QY 121 STKGCTACCTGGAGATCAGCTTCGGCTGCACCGCTTGTCTGCGGACGAGCGAGGCTTC 180
 Db 214 GTGGCTACCTGGAGATCAGCTTCGGCTGCACCGCTTGTCTGCGGACGAGCGAGGCTTC 273
 QY 181 TGCACCGCAACTGCGAGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCTAT 240
 Db 274 TGCACCGCAACTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTAT 333
 QY 241 CCCGAGCTGCGCTCTTTCGGGGGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCTAT 300
 Db 334 CCCGAGCTGCGCTCTTTCGGGGGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCTAT 393
 QY 301 CAGGCGCTGCGAGCTTCCGCGAGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTAT 360
 Db 394 CAGGCGCTGCGAGCTTCCGCGAGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTAT 453
 QY 361 CGCGCGAGCGCTTACAAAGTTCCTGCGAGTTCGCTTACTTCAAGGCAAAATATCTCCCGCAA 420
 Db 454 CGCGCGAGCGCTTACAAAGTTCCTGCGAGTTCGCTTACTTCAAGGCAAAATATCTCCCGCAA 513
 QY 421 GCGATGCGCGCTGCTCAGCTTTCCTGCTGAGGATCTTCTGATCAGAAATGATGAAGAGG 480
 Db 514 GCGATGCGCGCTGCTCAGCTTTCCTGCTGAGGATCTTCTGATCAGAAATGATGAAGAGG 573
 QY 481 AACATGGCATATATAAGAGCCTGCTGTCGCCGAGGACTACATTAAGACCTTGGAAAC 540
 Db 574 AACATGGCATATATAAGAGCCTGCTGTCGCCGAGGACTAC-TTAAAGACCTTGGAAAC 632
 QY 541 AAGTCATATAAGAGCTTGTTCATCGAGAGAGTGGGCGCATACAAACGGTGAGAACTGGAGA 600
 Db 633 AAGTCATATAAGAGCTTGTTCATCGAGAGAGTGGGCGCATACAAACGGTGAGAACTGGAGA 692
 QY 601 ACATCCATCACAGACATGAGCTGCCCTTCCGACTTCTTCAAGCTTTTACGAGTGT 660
 Db 693 ACATCCATCACAGACATGAGCTGCCCTTCCGACTTCTTCAAGCTTTTACGAGTGT 752
 QY 661 CTCGAGCGCTGCGAGGCTTCCAGGAGATCAAGGATTTCTCAAGGATTTCTACCTTTCCATA 720

Db 753 CTCGAGCCTGCGAGGGTTCAGGAGATCAAGGACTTCAGAGATTTCTACCTTTCCATA 812
 Qy 721 GCAGATCATATATAGTGAAGCTTTGGAATGCAAAATACAGTGTGAAGAGAACTTCACCCCA 780
 Db 813 GCAG----- 816
 Qy 781 GTTATAGGAGGTATCCGGTTTGAGAAATTTGTGGCTACCATGTATCTACTTGCAGTTT 840
 Db 817 ----- 816
 Qy 841 GCCTATTATAGTTGAACGACTTGAAGATGACAGCCCCCTCTGAGTCAGCTATCTCTC 900
 Db 817 -----TGAACGACCTGAGAAATGACAGCCCCCTCTGAGTCAGCTATCTCTC 863
 Qy 901 TTTGATCAGAAATGACAAGGTCTATGAGCAGAACCTTGTGTATTACCAAGTACCAAGGGAC 960
 Db 864 TTTGATCAGAAATGACAAGGTCTATGAGCAGAACCTTGTGTATTACCAAGTACCAAGGGAC 923
 Qy 961 ACTTGGGGCTCTCGATGAGCACTTCCAGCCAGACCTGAGCAGTTCAGTTCTTTAAT 1020
 Db 924 ACTTGGGGCTCTCGATGAGCACTTCCAGCCAGACCTGAGCAGTTCAGTTCTTTAAT 983
 Qy 1021 GTGACCACACTCCAGAGGAGCTGTATGACTTTTGTAAAGGAAATATATGATGATGAT 1080
 Db 984 GTGACCACACTCCAGAGGAGCTGTATGACTTTTGTAAAGGAAATATATGATGATGAT 1043
 Qy 1081 GAGGAGAGTGTGGAATATGTGATGACCTTTGGAACTGGAGGACCCAGC 1134
 Db 1044 GAGGAGAGTGTGGAATATGTGATGACCTTTGGAACTGGAGGACCCAGC 1097

RESULT 12
 BD148800
 LOCUS
 DEFINITION
 Primer for synthesizing full-length cDNA and use thereof.
 ACCESSION
 BD148800
 VERSION
 BD148800.1 GI:27854558
 KEYWORDS
 JP 2002191363-A/3643
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Oca, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
 Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
 Primer for synthesizing full-length cDNA and use thereof
 Patent: JP 2002191363-A 3643 09-JUL-2002;
 HELIX RESEARCH INSTITUTE
 OS Homo sapiens (human)
 PN JP 2002191363-A/3643
 PD 09-JUL-2002
 PF 28-JUL-2000 JP 2000280990
 PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
 PI SAITO,
 PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
 PI KEIICHI NAGAI, TETSUJI OTSUKI
 PC
 C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
 10,
 PC C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
 Primer for synthesizing full-length cDNA and use thereof FH Key
 Location/Qualifiers
 FT source 1. 818
 FT Location/Qualifiers
 1. 818
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 159 a 271 c 236 g 148 t
 ORIGIN
 Query Match 51.8%; Score 587.6; DB 6; Length 818;

Best Local Similarity 96.2%; Pred. No. 8.1e-89;
 Matches 631; Conservative 2; Mismatches 18; Indels 5; Gaps 3;
 Qy 1 GGGCGCCCCCAATACGAAGCTTACAGCTTCCGAGCTTCCACGGAGAGAGTGTATGCCG 60
 Db 109 GGGCGCCCCCAATACGAAGCTTACAGCTTCCGAGCTTCCACGGAGAGAGTGTATGCCG 168
 Qy 61 CTCAGTTCGGCTTACCGGACGCGCTGGAAGTACAGCGGCGAGCACTTGGGCGGAGAGC 120
 Db 169 CTCAGTTCGGCTTACCGGACGCGCTGGAAGTACAGCGGCGAGCACTTGGGCGGAGAGC 228
 Qy 121 STGGCTACCTGAGATCAGCTTCGGCTGACACCGCTTCTGCGGAGAGAGAGGAGGCTTC 180
 Db 229 GTGGCTACCTGAGATCAGCTTCGGCTGACACCGCTTCTGCGGAGAGAGAGGAGGCTTC 288
 Qy 181 TGGCACCGCAACTGACGCGCGCGCGAGCCCGCGGAGCGCGCGCGGCTTCGCCAGTAT 240
 Db 289 TGGCACCGCAACTGACGCGCGCGCGAGCCCGCGGAGCGCGCGCGGCTTCGCCAGTAT 348
 Qy 241 CCGAGCTGGCTTCTTGGGGGCTTCTGCGCGCGGCGCTGCTCAAGGCTGCAAG 300
 Db 349 CCGAGCTGGCTTCTTGGGGGCTTCTGCGCGCGGCGCTGCTCAAGGCTGCAAG 408
 Qy 301 CAGGGCTTCCAGCTTCCGCGAGTCCAGCCCGCGGAGCGCGGAGTCTGCGGAGCTTCAG 360
 Db 409 CAGGGCTTCCAGCTTCCGCGAGTCCAGCCCGCGGAGCGCGGAGTCTGCGGAGCTTCAG 468
 Qy 361 CGCGCGAGCGCTTACAGTTCCTGCGAGTTCGCTTCAAGGCAAAATATCTCCCAAA 420
 Db 469 CGCGCGAGCGCTTACAGTTCCTGCGAGTTCGCTTCAAGGCAAAATATCTCCCAAA 528
 Qy 421 GCCATCCGCGCTGCTCACACCTTCTTCTGAGCACTTCCTGATGAGCAAAATATGAGAGG 480
 Db 529 GCCATCCGCGCTGCTCACACCTTCTTCTGAGCACTTCCTGATGAGCAAAATATGAGAGG 588
 Qy 481 AACATGGCATATTATAGAGCTTCCCTGCGGAGGAGTCTTCTTAAAGACCTTGAAGACC 540
 Db 589 AACATGGCATATTATAGAGCTTCTGCGGAGGAGTCTTCTTAAAGACCTTGAAGACC 647
 Qy 541 AAGT--CATATGAAGCTTCTTCTGAGCACTTCGCGAGTTCGCGGAGTCTGAGAGTGA 598
 Db 648 AAGTTCATATGAAGGCTTCTTCTGAGCACTTCGCGAGTTCGCGGAGTCTGAGAGTGA 707
 Qy 599 GAAATTCATCACAGACATGAGCTGCGCC--TTCCCGAGCTTCTTCAAGACCTTTT 652
 Db 708 GAACATTCATCACAGACATGAGCTGCGCCCTTCCGAGTTCGCGAGTTCCTTCAAGAGCTTT 763

RESULT 13
 GGCARPR
 LOCUS
 DEFINITION
 G.gallus mRNA for cartilage associated protein.
 ACCESSION
 X97607
 VERSION
 X97607.1 GI:1296525
 KEYWORDS
 cartilage associated protein.
 SOURCE
 Gallus gallus (chicken)
 ORGANISM
 Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 Castagnola, P., Gennari, M., Morello, R., Tonachini, L., Marin, O.,
 Gaggero, A. and Cancedda, R.
 Cartilage associated protein (CASP) is a novel developmentally
 regulated chick embryo protein
 J. Cell. Sci. 110 (Pt 12), 1351-1359 (1997)
 97360293
 9217321
 REFERENCE 2 (bases 1 to 1665)
 AUTHORS
 Castagnola, P.
 TITLE
 Direct Submission
 Submitted (29-APR-1996) P. Castagnola, Centro di Biotecnologie
 Avanzate, Differenzamento Cellulare, L.go R. Benzi, 10 Genova,

```

16132, ITALY
FEATURES
  Location/Qualifiers
    1..1665
      /organism="Gallus gallus"
      /mol_type="mRNA"
      /db_xref="taxon:9031"
      /cell_type="chondrocyte"
      /lab_host="E.coli"
      /lab_host="E.coli"
    43..858
      /codon_start=1
      /product="cartilage associated protein"
      /protein_id="CA66206.1"
      /db_xref="GI:1296526"
      /translation="MMTLLAALLATGAGYRYSPFRPRLMPLSAYVYGLDQY
STEMPESVYLEVSLRYELLRDTAFCHNCSSAGPLTAPPADGALRLLAGV
LLRAQCDERKQGLPQAPQGRLELFFORPEYKYLQFAFYKANNLPKRAIAAHT
FLRHPPDERMQRMYKGIIPDAEHIKDLTKPYENLFVRAVAYNGDNWRKTSISD
MELALPDFFTYDDCIAACEGSEIKDFKDFLSIADHYIEVLACKVQFDIT"
    389 c 388 g 418 t
BASE COUNT
  470 a 389 c 388 g 418 t
ORIGIN
  Query Match 43.6%; Score 494.6; DB 5; Length 1665;
  Best Local Similarity 78.6%; Pred. No. 3e-73;
  Matches 603; Conservative 1; Mismatches 160; Indels 3; Gaps 1;
QY 1 GGGCGCGCCCAATACGACGCTACAGCTTCGGCAGCTTCCACGGGACGAGCTGATGCGG 60
DB 79 GCGGCGCGCGAGTACGAGCGGTACAGCTTCGGCAGCTTCCGCGGACGAGCTGATGCGG 138
QY 61 CTCGAGTCGGCTACCGCGAGCGCTCGACAGCTACGAGCGGCGAGCACTGGCGCGAGG 120
DB 139 CTGGAACTCCGCTACCGCTACCGCTCGGCTGACAGTACGAGCGGAGCACTGGCGCGAGG 198
QY 121 STKGCTACCTGGAGATACGCTCCGCTGACCGCTTGTGCGGACGAGCGAGCGCTTC 180
DB 199 GTCACTACCTGGAGTACGAGTACGCTTCGCTGCGGCTGCGGACGAGCGAGCGCTTC 258
QY 181 TGCCACCGCAACTGAG--CGCGCGCGCGAGCGCGCGCGCGCGCGCTCGCGAGC 237
DB 259 TGCCACCACTACGAGTCCGCGCGCGCTACCGCGCGCGCGCGCGCGCGCGCGCGCTG 318
QY 238 TATCCGAGCTGCGCTTCGCGGCGCTGCTGCGGCGCGCGCGCGCGCGCGCGCGCTG 297
DB 319 CTGCGCGAGCTGCGCTGCTGCGGCGAGTGTGCGGCGAGCGCAATGCTGCGGCGCTGC 378
QY 298 AAGCAGGCGCTGCGCGAGCTTCGCGCGAGTCCCGAGCGCGCGCGCGCGCGCGCGCTTC 357
DB 379 AAGCAGGCGTTCGCGCGCTTCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCTC 438
QY 358 CAGCGCGCGAGCGCTACAGTTCCTGCGAGTTCGCTTACTTCAAGCAATATCTCCCG 417
DB 439 CAGCGCGCGAGCGCTACAGTTCCTGCGAGTTCGCTTACTTCAAGCAATATCTCCCG 498
QY 418 AAAGCGCATCGCGCTGCTCACACCTTTCTACTGAAGCATCTCTGATGACGAATGATGAAG 477
DB 499 AAAGCTATTGAGCAGCTCACACATTTCTTGAACATCCAGATGATGAATGATGAAC 558
QY 478 AGCAACATGGCATATATAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 537
DB 559 AGAATATGGCTACTATAAGAGCATACCTGATCTGAGGAGCAGCATTAAGAGCTTGGAG 618
QY 538 ACCAAGTCATATGAAGCGCTTTCATCGAGCAGTGGCGGCATACCAACGCTGAGCACTGG 597
DB 619 ACAAGCGCTTACGAGAACTCTTTTGTGAGGCGCTGTGAGAGCATACAAATGCTGACACTGG 678
QY 598 AGCAACATCCATCAGCAGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 657
DB 679 AGAATATCACTCAGATATGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 738
QY 658 TGTCTCGAGCTGCGAGGTTCCAGGAGATCAGAGCTTCAAGAGCTTCAAGAGTTTCACTTCC 717

```

```

739 TGTATAGCAGCCTGTGAAGGCTCCCGAGAGATTAAAGATTAAAGACTTCTATCTCT 788
718 ATAGCAGATCATATGTAGAGTTCTGGAATGCAAAATACAGTGTGA 764
799 ATTGAGATCATATATTGAAGTCTCTGATGCAAAAGTGAGTTCGA 845

BC011701 1745 bp mRNA linear PRI 02-AUG-2001
Homo sapiens, nucleolar autoantigen (55KD) similar to rat
synaptonemal complex protein, clone MGC:19699 IMAGE:3510378, mRNA,
complete cds.
ACCESSION BC011701
VERSION BC011701.1 GI:15079791
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1745)
AUTHORS Strausberg, R.
TITLES Direct Submission
JOURNAL Submitted (30-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgabbs@mail.nih.gov
Tissue procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@nigri.nih.gov
Shenchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
Lim, M., Maduro, Q.L., Masello, C., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Snyder, B., Stantipop, S., Thomas, P.J.,
Tiongson, E.E., Touchman, J.W., Tsurgon, C., Vogt, J.L., Walker, M.A.,
Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL plate: 27 Row: d Column: 23
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.
FEATURES
  Location/Qualifiers
    1..1745
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="LocusID:10609"
      /db_xref="taxon:9606"
      /clone="MGC:19699 IMAGE:3510378"
      /tissue_type="Brain, neuroblastoma"
      /clone_lib="NIH MGC_19"
      /lab_host="DH10B-R"
      /note="Vector: pOTS7"
    26..1339
      /codon_start=1
      /product="nucleolar autoantigen (55KD) similar to rat
synaptonemal complex protein"
      /protein_id="AAH11701.1"
      /db_xref="GI:15079792"
      /translation="MARVANGILLMLLGSAGAOVEKYSPGFPEPDLMPAAVGHAL
SPGESWREARYLEALRLRLRUSEAFCHANCSPAPAKPDPDGGGRADENACE
LKLFRVLRACLRRCRTTAPQVTPPQLRDFQKLPYQILYALFKANRLER
AVAAATYFLQNPKELTAKYINYYQGNLDVADESLTDLAQPYEAVFLAVKLYNSG

```

DFRSTEDMERALSEVLAVFARCLAGCEGAHQVDKDFYPAIADLFAELSLQCKVDC
 ANLTPNVGIFYVDKFAVYHYLYQFAYYKLANDVROARSAASVYMLFPKDSVMQNLV
 YVHFHARGLBEEEDFPQREEMLYNQTAELRELLEFTHMYLQSDDEMBELTEPEPL
 EPDALSDAEFEGEGYBEGMVDWQSPDAKDEAEPEPELA"
 BASE COUNT 399 a 570 c 611 g 309 t
 ORIGIN

Query Match 40.2%; Score 455.4; DB 9; Length 1889;
 Best Local Similarity 55.1%; Pred. No. 1.1e-66;
 Matches 723; Conservative 2; Mismatches 368; Indels 18; Gaps 3;
 Qy 5 GCGCCCAATACGAACGCTACAGCTTCCGCGAGCTTCCACGGGACGAGCTGATCGCGCTCG 64
 Db 96 GGGCGCAGTACGAGAAGTACAGCTTCCGGGGCTTCCCGCCGAGGACCTGATCGCGCTGG 155
 Qy 65 AGTCGGCTACCGGCGAGCGCTGGACAGTACAGGGGCGAGCACTGGSCCGGAGGCGSTKG 124
 Db 156 CCGCGGCTACGGGCGAGCTTGGAGCAGTACAGGGAGAGCTGGCGCGAGAGCGCGC 215
 Qy 125 GCTACCTGGAGATCAGCCTCGCGCTCCACCGCTTGTGCGGACAGCGAGGACCTTCTGCC 184
 Db 216 GCTACCTGGAGGCGGCGCTCGCGCTGACCGGCTTCCGCGACAGCGAGGCGCTTCTGCC 275
 Qy 185 ACGCCAATCAGCG-----CGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCA 235
 Db 276 ACGCCAATCAGCG 335
 Qy 236 GCTA-----TCCGCGAGCTGGCGCTTTCGGGGGCGCTGTCGCGCGCGCGCGCGCGCTCA 289
 Db 336 ACAGTGGGCGCTGCGAGCTGGCGCTTTCGGCGCGCTGTCGCGCGCGCGCGCGCGCGCG 395
 Qy 290 AGCGCTGCAAGCAGGCGCTCCAGCGCTTCCGCGAGTCCCGCGCGCGCGCGCGCGCGCG 349
 Db 396 GCGCTGCAAGCAGCGCTGCGCGCTTCCAGGTGCGCTTCCCGCGCGCGCGCGCGCGCTGC 455
 Qy 350 CGGACTTCCAGCGCGCGCGCGCGCTTCCAGGTGCGCTTCCAGGTGCGCTTCCAGGTGCGCT 409
 Db 456 GTGACTTCCAGCGCGCGCTGCGCGCTTCCAGGTGCGCTTCCAGGTGCGCTTCCAGGTGCG 515
 Qy 410 ATCTCCCGCAAGCGCTGCGCGCTGCTCAGCGCTTCCAGGTGCGCTTCCAGGTGCGCTTCC 469
 Db 516 GGTGGAGAGGCGGCTGCGCGCGCTTCCAGGTGCGCTTCCAGGTGCGCTTCCAGGTGCGCT 575
 Qy 470 TGATGAAGAGGACATGGCAGATATTA---GAGCGCTGCTGCTGCGCGCGCGCGCGCGCTCA 526
 Db 576 TGACCGCAAGTATCAACTACTATCAGGGGATGCTGCGCGCGCGCGCGCGCGCGCGCGCTCA 635
 Qy 527 AAGACCTGGAAACCAAGTCAATATGAAGGCTGTTATCCGAGCAGTGGCGCGCGCGCGCGCTCA 586
 Db 636 CGGACCTAGAGGCGCGCGCTTCCAGGTGCGCTTCCAGGTGCGCTTCCAGGTGCGCTTCCAG 695
 Qy 587 GTGAGACTGGAGAAATCCATCAGAGATGGAGTGGCGCTTCCCGCGCGCGCGCGCGCTCAAG 646
 Db 696 GCGGGGATTTCCGCGAGGAGCGCGGAGCATGGAGGGGCGCTTGTGAGGACTCTGGCGAG 755
 Qy 647 CCTTTTACGAGTGTCTCGAGCGCTGCGAGGGTTTCCAGGGAGATCAAGGACTTCAAGGATT 706
 Db 756 TCTTTGCGCGCTGCTGCGCGCTGTAAGGGGCGCGCTGAGGAGGCGCGCTTCAAGGACT 815
 Qy 707 TCTACCTTCCATAGCAGATCATATGTAAGTCTGGAATGCAAAATACAGTGTGAG 766
 Db 816 TCTACCGCGCATAGCAGATCTCTTTCGAGAGTCCCTGCGAGTCAAGGTGAGTGTGAGG 875
 Qy 767 AGAACCTTCAACCGAGTTATAGGAGGCTATCCGGTTGAGAAATTTGTGGCTACCATGTATC 826
 Db 876 CCAATTTGACCCCGCAATGTGGTGGCTACTTCTGCGCAAGTTCTGCGCGCGCGCGCGCTG 935
 Qy 827 ATTACTTGCAGTTTGCCTATTATAGTTGAAGCGAGCTGAAGATGAGCGCGCGCGCGCTG 886
 Db 936 ACTACCTGCGAGTTTGCCTACTATAGTTGATGATGTCGCGAGGCTGCGCGCGCGCGCG 995
 Qy 887 TCAGCTATCTGCTCTTTGATCAGAAATGACAAAGTCAAGGATGAGCGAGAACTGGGTATTACC 946

Search completed: January 30, 2004, 14:49:21
 Job time : 4134.45 secs

Db 996 CCAGCTACATGCTCTTCGACCCCAAGGACAGCGCTCATGAGCAGAACCTGGTGATATTACC 1055
 Qy 947 AGTACCAACAGGGACACTTGGGGCGCTCTCGGATGAGCACTTCCAGCGCGCGAGCCTGAAAGCAG 1006
 Db 1056 GGTTCACCGGGCTCGCTGGGGCTTGAAGAGGAGGACTTCCAGCGCGCGAGGAGGCCA 1115
 Qy 1007 TTCAGTTCTTTAATGTGACCACACTCCAGAAAGGAGCTGTATGACTTTGCTAAGGAAATA 1066
 Db 1116 TGCTCTACCAACAGACCGCGAGCTGCGGAGCTTCTGGAGTTTACCCACATGTATACC 1175
 Qy 1067 TAATGATGATGATGAGGAGAAAGTTTGTGA 1097
 Db 1176 TGCACTCAGATGATGAGATGAGCTGGAGGA 1206